

# Riemannian Calculus of Variations using Strongly Typed Tensor Calculus

Victor Dods

2012.12.11

## Abstract

In this paper, the notion of strongly typed language will be borrowed from the field of computer programming to introduce a calculational framework for linear algebra and tensor calculus for the purpose of detecting errors resulting from inherent misuse of objects and for finding natural formulations of various objects. A tensor bundle formalism, crucially relying on the notion of pullback bundle, will be used to create a rich type system with which to distinguish objects. The type system and relevant notation is designed to “telescope” to accomodate a level of detail appropriate to a set of calculations. Various techniques using this formalism will be developed and demonstrated with the goal of providing a relatively complete and uniform method of coordinate-free computation.

The calculus of variations pertaining to maps between Riemannian manifolds will be formulated using the strongly typed tensor formalism and associated techniques. Energy functionals defined in terms of first order Lagrangians are the focus of the second half of this paper, in which the first variation, the Euler-Lagrange equations, and the second variation of such functionals will be derived.

## Introduction

Many important differential equations have a variational origin, being derived as the Euler-Lagrange equations for a particular functional on some space of functions. The variational approach lends itself particularly to physics, in which conservation of energy or minimization of action is a central concept. The naturality of such formulations can’t be understated, as solutions to such problems often depend critically on the inherent geometry of the underlying objects. For example, solutions to Laplace’s equation for a real valued function (e.g. modeling steady-state heat flow) on a Riemannian manifold depend qualitatively on the topology of the manifold (e.g. harmonic functions on a closed Riemannian manifold are necessarily constant, which makes sense geometrically because there is no boundary through which heat can escape).

A central concept in the field of software design is that of **information hiding** [16], in which a computer program is organized into modules, each presenting an abstract public interface. Other parts of the program can interact only through the presented interface, and the details of how each module works are hidden, thereby preventing interference in the implementation details which are not required by the inherent structure of the module. This concept has clear usefulness in the field of mathematics as well. For example, there are several formulations of the real numbers (e.g. equivalence classes of Cauchy sequences of rational numbers, Dedekind cuts, decimal expansions, etc), but their particulars are instances of what are known as **implementation details**, and the details of each particular implementation are irrelevant in most areas of mathematics, which only use the inherent properties of the real numbers as a complete, totally ordered field. Of course, at certain levels, it is useful or necessary to “open up the box” [go past the public interface] and work with a particular representation of the real numbers.

Information hiding is characteristic of abstract mathematics, in which general results are proved about abstract mathematical objects without using any particular implementation of said objects. These results can then be used modularly in other proofs, just as the functionality of a computer program is organized into modularized objects and functions. For example, a fixed point theorem for contractive mappings on closed

sets in Banach spaces, but a particular application of this theorem renders an existence and uniqueness theorem for first order ODEs [19, pgs. 59, 62].

A loose conceptual analogy for modularity is that of diagonalizing a linear operator. A basis of eigenvectors are chosen so that the action of the linear operator on each eigenspace has a particularly simple expression, and distinct eigenspaces do not interact with respect to the operator’s action. In this analogy, the eigenvectors then correspond to individual lemmas, and the linear operator corresponds to a large theorem which uses each lemma. Decomposing the proof of the main result in terms of non-interacting lemmas simplifies the proof considerably, just as it simplifies the quantification of the linear operator. The term “orthogonal” has been borrowed by software design to describe two program modules whose functionality is independent [18, Chapter 4, Section 2]. Orthogonality in software design is highly desirable as it generally eases program implementation and program correctness verification, as the human designers are only capable of keeping track of a certain finite number of details simultaneously [13]. The scope of each detail level of the design is limited in complexity, making the overall design easier to comprehend.

This technique in software design carries over directly to proof design, where it is desirable (elegant) to write proofs and do calculations without introducing extraneous details, such as choice of bases in vector spaces or local coordinates in manifolds. Because such choices are generally non-unique, they can often obscure the inherent structure of the relevant objects by introducing artifacts arising from properties of the particular details used to implement said objects. For example, the choice of a particular local coordinate chart on a manifold artificially imposes an additive structure on a neighborhood of the manifold, but such a structure has nothing to do with the inherent geometry of the manifold. Furthermore, the descent to this “lower level” of calculation discards some type information, representing points in a manifold as Euclidean vectors, thereby losing the ability to distinguish points from different manifolds, or even different localities in the same manifold.

This paper makes a particular emphasis on natural formulations and calculations in order to expose the underlying geometric structures rather than relying on coordinate-based expressions. The construction of the “full” direct sum and “full” tensor product bundles are used in combination with induced covariant derivatives to this end.

## Notation and Conventions

Let all vector spaces, manifolds and [fiber] bundles be real and finite-dimensional unless otherwise noted (this allows the canonical identification  $V^{**} \cong V$  for a vector space or vector bundle  $V$ ), and let all tensor products be over  $\mathbb{R}$ . The unqualified term “bundle” will mean “fiber bundle”. The Einstein summation convention will be assumed in situations when indexed tensors are used for computation.

Unary operators are understood to have higher binding precedence than binary operators, and super and subscripts are understood to have the highest binding precedence. For example, the expression  $\nabla X_{,M} \circ \phi$  would be parenthesized as  $(\nabla(X_{,M})) \circ \phi$ .

Apart from the obvious purpose of providing a concise and central reference for the notation in this paper, the following notation index serves to illustrate the use of telescoping notation (see Section 2). The high-level (terse notation which requires the reader to do more work in type inference but is more agile), mid-level, and low-level (completely type-specified, requiring little work on the part of the reader) notations are presented side-by-side with their definitions.

Let  $I \subseteq \mathbb{R}$  be a neighborhood of 0, let  $\epsilon, i$  each be coordinates on  $I$ , let  $A, A_1, \dots, A_n, B$  be sets, let  $M, N$  be manifolds, let  $\phi \in C^\infty(M, N)$ , let  $\pi_M^A: A \rightarrow M$  and  $\pi_N^H: H \rightarrow N$  and be vector bundles, where  $A = E, F, F_1, \dots, F_n, G$ , let  $U, V, V_1, \dots, V_n, W$  be vector spaces, and let  $c_i \in \Gamma(F_i \otimes_M T^*M)$  such that

$$c_1 \oplus_M \dots \oplus_M c_n \in \Gamma((F_1 \oplus_M \dots \oplus_M F_n) \otimes_M T^*M)$$

is a vector bundle isomorphism.

High-	Mid-	Low-level	Description
Variations; variational derivatives; tangent vectors.			
$m_\epsilon$ $\delta$ $\delta m_\epsilon$	$m$ $\delta_\epsilon$ $\delta_\epsilon m$	Variation of a point in $M$ ; $I \ni \epsilon \mapsto m_\epsilon \in M$ ; $m: I \rightarrow M$ . Variational derivative; $\delta_\epsilon := \frac{\partial}{\partial \epsilon} \big _{\epsilon=0}$ . Tangent vector; linearization of a variation; $\delta m_\epsilon \in T_{m_0} M$ ; $\delta_\epsilon m \in T_{m(0)} M$ .	
Projection maps; canonical isomorphisms; bundle-related maps and spaces.			
pr	$\text{pr}_i$ $\text{pr}_{A_i}$	$\text{pr}_i^{A_1 \times \cdots \times A_n}$ $\text{pr}_{A_i}^{A_1 \times \cdots \times A_n}$	Set-theoretic projection onto $i$ th factor or named factor; $\text{pr}_i^{A_1 \times \cdots \times A_n}: A_1 \times \cdots \times A_n \rightarrow A_i$ .
$\iota$ $\pi$ $\rho$	$\iota_B, \iota^A$ $\pi_M, \pi^F$ $\rho_H, \rho^{\phi^* H}$	$\iota_B^A$ $\pi_M^F$ $\rho_H^{\phi^* H}$	Canonical isomorphism; $\iota_B^A: A \rightarrow B$ ; $\iota_A^B := (\iota_B^A)^{-1}$ . Bundle projection map; $\pi_M^F: F \rightarrow M$ . Pullback bundle fiber projection map; $\rho_H^{\phi^* H}: \phi^* H \rightarrow H$ .
Trivial bundle constructions and projection maps.			
$M \times N \rightarrow N$  $M \times N \rightarrow M$	$M \rtimes N \rightarrow N$  $M \rtimes N \rightarrow M$	Trivial bundle over $N$ ; $M \rtimes N := M \times N$ ; $\pi_N^{M \rtimes N}: M \rtimes N \rightarrow N, (m, n) \mapsto n$ . Trivial bundle over $M$ ; $M \rtimes N := M \times N$ ; $\pi_M^{M \rtimes N}: M \rtimes N \rightarrow M, (m, n) \mapsto m$ .	
Shared base-space bundle constructions and projection maps.			
$E \times F \rightarrow M$  $E \oplus F \rightarrow M$  $E \otimes F \rightarrow M$	$E \times_M F \rightarrow M$  $E \oplus_M F \rightarrow M$  $E \otimes_M F \rightarrow M$	Direct product; $E \times_M F := \coprod_{m \in M} E_m \times F_m$ ; $\pi_M^{E \times_M F}(e, f) := \pi_M^E(e) \equiv \pi_M^F(f)$ . Whitney sum; $E \oplus_M F := \coprod_{m \in M} E_m \oplus F_m$ ; $\pi_M^{E \oplus_M F}(e \oplus f) := \pi_M^E(e) \equiv \pi_M^F(f)$ . Tensor product; $E \otimes_M F := \coprod_{m \in M} E_m \otimes F_m$ ; $\pi_M^{E \otimes_M F}(c^{ij} e_i \otimes f_j) := \pi_M^E(e_k) \equiv \pi_M^F(f_\ell)$ (for any $k, \ell$ ).	
Separate base-space bundle constructions and projection maps.			
$E \times H \rightarrow M \times N$  $E \oplus H \rightarrow M \times N$  $E \otimes H \rightarrow M \times N$	$E \times_{M \times N} H \rightarrow M \times N$  $E \oplus_{M \times N} H \rightarrow M \times N$  $E \otimes_{M \times N} H \rightarrow M \times N$	Direct product; $E \times_{M \times N} H := \coprod_{(m, n) \in M \times N} E_m \times H_n$ . $\pi_{M \times N}^{E \times_{M \times N} H}(e, h) := (\pi_M^E(e), \pi_N^H(h))$ . Whitney sum; $E \oplus_{M \times N} H := \coprod_{(m, n) \in M \times N} E_m \oplus H_n$ . $\pi_{M \times N}^{E \oplus_{M \times N} H}(e \oplus h) := (\pi_M^E(e), \pi_N^H(h))$ . Tensor product; $E \otimes_{M \times N} H := \coprod_{(m, n) \in M \times N} E_m \otimes H_n$ . $\pi_{M \times N}^{E \otimes_{M \times N} H}(c^{ij} e_i \otimes h_j) := (\pi_M^E(e_k), \pi_N^H(h_\ell))$ (for any $k, \ell$ ).	
Trace; natural pairing; tensor/tensor field contraction. Simple tensor expressions are extended linearly.			
tr $\alpha \cdot v$ $A \cdot B$  $S \cdot^n T$ $S: T, S \cdot^2 T$	$\text{tr}_V$ $\alpha \cdot_V v$ $A \cdot_V B$  $S \cdot_{V_1 \otimes \cdots \otimes V_n} T$ $S \cdot_{V_1 \otimes V_2} T$	Trace on $V$ ; $\text{tr}_V: V^* \otimes V \rightarrow \mathbb{R}, \alpha \otimes v \mapsto \alpha(v)$ . Natural pairing; $\cdot_V: V^* \times V \rightarrow \mathbb{R}, (\alpha, v) \mapsto \alpha(v)$ . Tensor contraction; $\cdot_V: (U \otimes V^*) \times (V \otimes W) \rightarrow U \otimes W$ , $(u \otimes \alpha) \cdot_V (v \otimes w) := u \otimes (\alpha \cdot_V v) \otimes w \equiv \alpha(v) u \otimes w$ . Alternate for $\cdot_V$ , where $V = V_1 \otimes \cdots \otimes V_n$ . Special notation for $n = 2$ .	
tr $\sigma \cdot f$ $A \cdot f$ $S \cdot T$  $S \cdot^n T$ $S: T, S \cdot^2 T$	$\text{tr}_F$ $\sigma \cdot_F f$ $A \cdot_F f$ $S \cdot_F T$  $S \cdot_{F_1 \otimes_M \cdots \otimes_M F_n} T$ $S \cdot_{F_1 \otimes_M F_2} T$	Trace on $F \rightarrow M$ ; $\text{tr}_F: \Gamma(F^* \otimes_M F) \rightarrow C^\infty(M, \mathbb{R})$ , $[\text{tr}_F(\sigma \otimes_M f)](m) := \sigma(m) \cdot_{F_m} f(m)$ for $m \in M$ . Natural pairing; $\cdot_F: \Gamma(F^*) \times \Gamma(F) \rightarrow C^\infty(M, \mathbb{R})$ , $(\sigma \cdot_F f)(m) := \sigma(m) \cdot_{F_m} f(m)$ for $m \in M$ . Natural pairing; $\cdot_F: \Gamma(E \otimes_M F^*) \times F \rightarrow E$ , $(e \otimes_M \sigma) \cdot_F f := e(m) (\sigma(m) \cdot_{F_m} f) \in E_m$ ; $m := \pi_M^F(f)$ . Tensor field contraction; pointwise tensor contraction; $\cdot_F: \Gamma(E \otimes_M F^*) \times \Gamma(F \otimes_M G) \rightarrow \Gamma(E \otimes_M G)$ , $[(e \otimes \sigma) \cdot_F (f \otimes g)](m) := (\sigma(m) \cdot_{F_m} f(m)) e(m) \otimes g(m)$ . Alternate for $\cdot_F$ , where $F = F_1 \otimes_M \cdots \otimes_M F_n$ . Special notation for $n = 2$ .	

High-	Mid-	Low-level	Description
Permutations of tensors and tensor fields.			
$A^\sigma, A \cdot^n \sigma$		$A \cdot V_1^* \otimes \dots \otimes V_n^* \sigma$	Right-action of permutations on $n$ -tensors/ $n$ -tensor fields; $(v_1 \otimes \dots \otimes v_n)^\sigma := v_{\sigma^{-1}(1)} \otimes \dots \otimes v_{\sigma^{-1}(n)}$ ; $(A^\sigma)^\tau = A^{\sigma\tau}$ .
Spaces of sections of bundles.			
$\Gamma(H), \Gamma(\pi_N^H)$			Space of smooth sections of the bundle $\pi_N^H$ ;
			$\Gamma(H) := \{h \in C^\infty(N, H) \mid \pi_N^H \circ h = \text{Id}_N\}$ .
$\Gamma_\phi(H), \Gamma_\phi(\pi_N^H)$			Space of smooth sections of $\pi_N^H$ along $\phi$ ;
			$\Gamma_\phi(H) := \{h \in C^\infty(M, H) \mid \pi_N^H \circ h = \phi\}$ .
Vertical bundle, pullback bundle, projection maps, pullback of sections.			
$VE \rightarrow E$			Vertical bundle over $E \rightarrow M$ ; $VE := \ker T\pi_M^E \leq TE$ . projection map $\pi_E^{VE} := \pi_E^{TE} _{VE}$ .
$\phi^*H \rightarrow M$			Pullback bundle; $\phi^*H := \{(m, h) \in M \times H \mid \phi(m) = \pi(h)\}$ . $\pi_M^{\phi^*H}(m, h) := m$ ; $\rho_H^{\phi^*H}(m, h) := h$ .
$\phi^*h$			Pullback of section $h \in \Gamma(H)$ ; $\phi^*h \in \Gamma(\phi^*H)$ defined by $\rho_H^{\phi^*H} \circ \phi^*h = h$ ; $h \in \Gamma(H)$ .
Covariant derivatives; partial covariant derivatives.			
$\nabla L$	$\nabla L$ $\nabla^{M \rightarrow \mathbb{R}} L$	$\nabla^{M \rightarrow \mathbb{R}} L$	Natural linear covariant derivative; differential of functions; $\nabla^{M \rightarrow \mathbb{R}} L := dL \in \Gamma(T^*M)$ , where $L \in C^\infty(M, \mathbb{R})$ .
$\nabla X$	$\nabla X$ $\nabla^E X$	$\nabla^E X$	Linear covariant derivative on vector bundle $E \rightarrow M$ ; $\nabla^E X \in \Gamma(E \otimes_M T^*M)$ , where $X \in \Gamma(E)$ .
$\nabla \phi$	$\nabla \phi$ $\nabla^{M \rightarrow N} \phi$	$\nabla^{M \rightarrow N} \phi$	Tangent map as tensor field; $\nabla^{M \rightarrow N} \phi \in \Gamma(\phi^*TN \otimes_M T^*M)$ , where $\phi \in C^\infty(M, N)$ .
$\nabla \sigma$	$\nabla \phi \sigma$	$\nabla^{\phi^*H} \sigma$	Pullback covariant derivative; $\sigma \in \Gamma(\phi^*H)$ ; defined by $\nabla^{\phi^*H} \phi^*h = \phi^* \nabla^H h \cdot \phi^*_{TN} \nabla^{M \rightarrow N} \phi$ ; $h \in \Gamma(H)$ .
$L_{,c_1}, \dots, L_{,c_n}$  $X_{,c_1}, \dots, X_{,c_n}$  $\phi_{,M_1}, \dots, \phi_{,M_n}$			Partial differential of functions; $L_{,c_i} \in \Gamma(F_i^*)$ , defined by $\nabla^{M \rightarrow \mathbb{R}} L = \sum_{i=1}^n L_{,c_i} \cdot F_i c_i$ . Partial linear covariant derivative; $X_{,c_i} \in \Gamma(E \otimes_M F_i^*)$ , defined by $\nabla^E X = \sum_{i=1}^n X_{,c_i} \cdot F_i c_i$ . Partial derivative decomposition of tangent map; $\phi_{,M_i} \in \Gamma(\phi^*TN \otimes_M \text{pr}_i^* T^*M_i)$ , where $M = M_1 \times \dots \times M_n$ , $\text{pr}_i := \text{pr}_i^M$ , and $\nabla^{M \rightarrow N} \phi = \sum_{i=1}^n \phi_{,M_i} \cdot \text{pr}_i^* T M_i \nabla^{M \rightarrow M_i} \text{pr}_i$ .
Covariant Hessians.			
$\nabla^2 L$ $\nabla \nabla L$ $\nabla^2 X$ $\nabla \nabla X$ $\nabla^2 \phi$ $\nabla \nabla \phi$	$\nabla^{T^*M} \nabla^{M \rightarrow \mathbb{R}} L$ $\nabla^{T^*M} \nabla^{M \rightarrow \mathbb{R}} L$ $\nabla^{E \otimes T^*M} \nabla^E X$ $\nabla^{E \otimes_M T^*M} \nabla^E X$ $\nabla^{\phi^*TN \otimes T^*M} \nabla^{M \rightarrow N} \phi$ $\nabla^{\phi^*TN \otimes_M T^*M} \nabla^{M \rightarrow N} \phi$		Covariant Hessian of functions; $\nabla^2 L \in \Gamma(T^*M \otimes T^*M)$ ; $L \in C^\infty(M, \mathbb{R})$ . Covariant Hessian on vector bundle $E \rightarrow M$ ; $\nabla^2 X \in \Gamma(E \otimes T^*M \otimes T^*M)$ ; $X \in \Gamma(E)$ . Covariant Hessian of maps; $\nabla^2 \phi \in \Gamma(\phi^*TN \otimes T^*M \otimes T^*M)$ . $\phi \in C^\infty(M, N)$ .
Derivative conventions.			
$\nabla_X e$ $\nabla^n e \cdot^n (X_1 \otimes \dots \otimes X_{n-1} \otimes X_n)$  $R(X, Y) := -\nabla_X \nabla_Y + \nabla_Y \nabla_X + \nabla_{[X, Y]}$  $z: M \rightarrow M \times I, m \mapsto (m, 0)$ $z^* \partial_i = \delta_i$			Directional derivative notation; $\nabla_X e := \nabla e \cdot T M X$ . Iterated covariant derivative convention; defined by $(\nabla_{X_n} \nabla^{n-1} e) \cdot^{n-1} (X_1 \otimes \dots \otimes X_{n-1})$ . Curvature operator; $R(X, Y) e = \nabla^2 e : (X \otimes Y - Y \otimes X)$ .  Evaluation-at-zero map. Pullback formulation of derivative-at-zero.

For more on relevant introductory theory on manifolds, bundles and Riemannian geometry, see [10], [8], [12], [9].

## Part I

# Mathematical Setting

## 1 Using Strong Typing to Error-Check Calculations

Linear algebra is an excellent setting for discussion of the **strong typing** [1] of a language, a concept used in the design of computer programming languages. The idea is that when the human-readable source code of a program is compiled (translated into machine-readable instructions), the compiler (the program which performs this translation) or runtime (the software which executes the code) verifies that the program objects are being used in a well-defined way, producing an error for each operation that is not well-defined. For example, a vector-type value would not be allowed to be added to a permutation-type value, even though tuples of unsigned integers (i.e. bytes) are used by the computer to represent both, and the computer's processing unit could add together their byte-valued representations. However, such an operation would be meaningless with respect to the types of the operands. The result of the operation would depend on the non-canonical choice of representation for each object. Strong type checking has the advantage of catching many programming errors, including most importantly those resulting from an inherent misuse of the program's objects. Within this paper, certain type-explicit notations will be used to provide forms of type awareness conducive to error-checking.

An important example of semi-strong typing in math is Penrose's abstract index notation [17], modeled on Einstein's summation convention, in which linear algebra and tensor calculus are implemented using indexed objects (tensors) having a certain number and order of “up” and “down” indices (an abstraction of the genuine basis/coordinate expressions in which the indexed objects are arrays of scalars/functions). A non-indexed tensor is a scalar value, a tensor having a single up or down index is a vector or covector value respectively, a tensor having an up and a down index is an endomorphism, and so forth. The tensors are contracted by pairing a certain number of up indices with the same number of down indices, resulting in an object having as indices the uncontracted indices.

For example, given a finite-dimensional inner product space  $(V, g)$ , where  $g$  is a  $\binom{0}{2}$ -tensor (having the form  $g_{ij}$ , i.e. two down indices), a vector  $v \in V$  is a  $\binom{1}{0}$ -tensor, and the length of  $v$  is  $\sqrt{v^i g_{ij} v^j}$ . If  $\dim V > 1$ , then  $\bigwedge^2 V$  has positive dimension, its vectors each being  $\binom{2}{0}$ -tensors, and  $G_{ijkl} := g_{ik}g_{jl} - g_{il}g_{jk}$  is an inner product on  $\bigwedge^2 V$  (which must be a  $\binom{0}{4}$ -tensor in order to contract with two  $\binom{2}{0}$ -tensors).

Certain type errors are detected by use of abstract index notation in the form of index mismatch. For example, with  $(V, g)$  as above, if  $\alpha \in V^*$ , then  $\alpha$  is a  $\binom{0}{1}$ -tensor. Because of the repeated  $j$  down indices, the expression  $g_{ij}\alpha_j$  typically indicates a type error;  $g_{ij}$  can't contract with  $\alpha_j$  because of incompatible valence (valence being the number of up and down indices). Furthermore, multiplying a  $\binom{0}{2}$ -tensor with a  $\binom{0}{1}$ -tensor without contraction should result in a  $\binom{0}{3}$ -tensor, which should be denoted using three indices, as in  $g_{ij}\alpha_k$ .

The only explicit type information provided by abstract index notation is that of valence. The “semi” qualifier mentioned earlier is earned by the lack of distinction between the different spaces in which the tensors reside. For example, if  $U, V, W$  are finite-dimensional vector spaces, then linear maps  $A: U \rightarrow V$  and  $B: V \rightarrow W$  can be written as  $\binom{1}{1}$ -tensors, and their composition  $B \circ A: U \rightarrow W$  is written as the tensor contraction  $(B \circ A)_j^i = B_k^i A_j^k$ . However, while the expression  $A_k^i B_j^k$  makes sense in terms of valence compatibility (i.e. grammatically), the composition “ $A \circ B$ ” that it should represent is not well-defined. Thus this form of type error is not caught by abstract index notation, since the domains/codomains of the linear maps must be checked separately.

The use of dimensional analysis (the abstract use of units such as kilograms, seconds, etc) in Physics is an important precedent of strong typing. Each quantity has an associated “dimension” (this is a different meaning from the “dimension” of linear algebra) which is expressed as a fraction of powers of formal symbols. The ordinary algebraic rules for fractions and formal symbols are used for the dimensional components, with

the further requirement that addition and equality may only occur between quantities having the same dimension.

For example, if  $E$ ,  $M$  and  $C$  represent the dimensions of energy, mass and cost, respectively, and if the energy storage density  $\rho E/M$  of a battery manufacturing process is known (having dimensions energy per mass) and the manufacturing weight yield  $w M/C$  of the battery is known (having dimensions mass per cost), then under the algebraic conventions of dimensional analysis, calculating the energy storage per cost (which should have dimensions energy per cost) is simple;

$$\left(\rho \frac{E}{M}\right) \left(w \frac{M}{C}\right) = \rho w \frac{EM}{MC} = \rho w \frac{E}{C}$$

(the  $M$  symbols cancel in the fraction). Here, both  $\rho$  and  $w$  are real numbers, and besides using the well-definedness of real multiplication, no type-checking is done in the expression  $\rho w$ .

A contrasting example is the quantity  $\rho/w$ , having dimensions  $EC/M^2$ . However, these dimensions may be considered to be meaningless in the given context. The quantity's type adds meaning to the real-valued quantity, and while the quantity is well-defined as a real number, the uselessness of the type may indicate that an error has been made in the calculations. For example, a type mismatch between the two sides of an equation is a strong indication of error.

This is also a convenient way to think about the chain rule of calculus. If  $z(y)$ ,  $y(x)$ , and  $x$  measure real-valued quantities, then  $z(y(x))$  measures the quantity  $z$  with respect to quantity  $x$ . Using  $Z$ ,  $Y$  and  $X$  for the dimensions of the quantities  $z$ ,  $y$  and  $x$  respectively, the derivative  $\frac{dz}{dx}$  has units  $Z/X$ . When worked out, the dimensions for the quantities on either side of the equation  $\frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx}$  will match exactly, having a non-coincidental similarity to the calculation in the battery product example.

## 2 Telescoping Notation (aka Don't Fear the Verbosity)

Many of the computations developed in this paper will appear to be overly pedantic, owing to the decoration-heavy notation that will be introduced in Section 3. This decoration is largely for the purpose of tracking the myriad of types in the type system and to assist the human reader or writer in making sense of and error-checking the expressions involved. The pedantry in this paper plays the role of introducing the technique. The notation is designed to **telescope**<sup>1</sup>, meaning that there is a spectrum of notational decoration; from

- pedantically type-specified, verbose, and decoration-heavy, where [almost] no types must be inferred from context and there is little work or expertise required on the part of the reader, to
- somewhat decorated but more compact, where the reader must do a little bit of thinking to infer some types, all the way to
- tersely notated with minimal type decoration, where [almost] all types must be inferred from context and the reader must either do a lot of thinking or be relatively experienced.

Additionally, some of the chosen symbols are meant to obey the same telescoping range of specificity. For example, compare  $n$ -fold tensor contraction  $\cdot^n$  with type-specified  $\cdot_{V_1 \otimes \dots \otimes V_n}$  as discussed in Section 3, or the symbols  $\nabla$ ,  $\nabla$ , and  $\nabla$  as discussed in Section 10. Tersely notated computations can be seen in Section 10, while fully-verbose computations abound in the careful exposition of Part II.

## 3 Strongly-Typed Linear Algebra via Tensor Products

A fully strongly typed formulation of linear algebra will now be developed which enjoys a level of abstraction and flexibility similar to that of Penrose's abstract index notation. Emphasis will be placed on notational

<sup>1</sup>Credit for the notion of telescoping notation is due in part to David DeConde, during one of many enjoyable and insightful conversations.

and conceptual regularity via a tensor formalism, coupled with a notion of “untangled” expression which exploits and notationally depicts the associativity of linear composition.

If  $V$  denotes a finite-dimensional vector space, then let

$$\cdot_V: V^* \times V \rightarrow \mathbb{R}, (\alpha, v) \mapsto \alpha(v)$$

denote the **natural pairing** on  $V$ , and denote  $\cdot_V(\alpha, v)$  using the infix notation  $\alpha \cdot_V v$ . The natural pairing is a nondegenerate bilinear form and its bilinearity gives the expression  $\alpha \cdot_V v$  multiplicative semantics (distributivity and commutativity with scalar multiplication), thereby justifying the use of the infix  $\cdot$  operator normally reserved for multiplication. The natural pairing subscript  $V$  is seemingly pedantic, but will prove to be an invaluable tool for articulating and navigating the rich type system of the linear algebraic and vector bundle constructions used in this paper. When clear from context, the subscript  $V$  may be omitted.

Because  $V$  is finite-dimensional, it is reflexive (i.e. the canonical injection  $V \rightarrow V^{**}$ ,  $v \mapsto (\alpha \mapsto \alpha(v))$  is a linear isomorphism). Thus the natural pairing  $\cdot_{V^*}$  on  $V^*$  can be written naturally as

$$\cdot_{V^*}: V \times V^* \rightarrow \mathbb{R}, (v, \alpha) \mapsto \alpha(v).$$

Note that  $\alpha \cdot_V v = v \cdot_{V^*} \alpha$ . Though subtle, the distinction between  $\cdot_V$  and  $\cdot_{V^*}$  is important within the type system used in this paper.

Through a universal mapping property of multilinear maps, the bilinear forms  $\cdot_V$  and  $\cdot_{V^*}$  descend to the **natural trace** maps

$$\begin{aligned} \text{tr}_V: V^* \otimes V &\rightarrow \mathbb{R}, \alpha \otimes v \mapsto \alpha(v), \text{ and} \\ \text{tr}_{V^*}: V \otimes V^* &\rightarrow \mathbb{R}, v \otimes \alpha \mapsto \alpha(v), \end{aligned}$$

each extended linearly to non-simple tensors. These operations can also be called **tensor contraction**. Noting that  $(V^* \otimes V)^*$  and  $(V \otimes V^*)^*$  are canonically isomorphic to  $V \otimes V^*$  and  $V^* \otimes V$  respectively, then for each  $A \in V^* \otimes V$  and  $B \in V \otimes V^*$ , it follows that  $\text{tr}_V(A) = \text{Id}_{V^*} \cdot_{V^* \otimes V} A$  and  $\text{tr}_{V^*}(B) = \text{Id}_V \cdot_{V \otimes V^*} B$ .

**Definition 3.1** (Linear maps as tensors). Let  $V$  and  $W$  be finite-dimensional vector spaces, and let  $\text{Hom}(V, W)$  denote the space of vector space morphisms from  $V$  to  $W$  (i.e. linear maps). The linear isomorphism

$$\begin{aligned} W \otimes V^* &\rightarrow \text{Hom}(V, W), \\ w \otimes \alpha &\mapsto (V \rightarrow W, v \mapsto w(\alpha \cdot_V v)) \end{aligned}$$

(extended linearly to general tensors) will play a central conceptual role in the calculations employed in this paper, as it will facilitate constructions which would otherwise be awkward or difficult to express. Linear maps and appropriately typed tensor products will be identified via this isomorphism.

Given bases  $v_1, \dots, v_m \in V$  and  $w_1, \dots, w_n \in W$ , and dual bases  $v^1, \dots, v^m \in V^*$  and  $w^1, \dots, w^n \in W^*$ , a linear map  $A: V \rightarrow W$  can be written under the identification in (3.1) as

$$A = A_j^i w_i \otimes v^j,$$

where  $A_j^i = w^i \cdot_W A \cdot_V v_j \in \mathbb{R}$ , and in fact  $[A_j^i] \in M_{n \times m}(\mathbb{R})$  is the matrix representation of  $A$  with respect to the bases  $v_1, \dots, v_m \in V$  and  $w_1, \dots, w_n \in W$ , noting that the  $i$  and  $j$  indices denote the “output” and “input” components of  $A$  respectively. Tensors are therefore the strongly typed analog of matrices, where the  $W \otimes V^*$  type information is carried by the  $w_i \otimes v^j$  component.

One clarifying example of the tensor formulation is the adjoint operation of the natural pairing, also known as forming the dual of a linear map. It is straightforward to show that

$$\begin{aligned} *: W \otimes V^* &\rightarrow V^* \otimes W, \\ w \otimes \alpha &\mapsto \alpha \otimes w, \end{aligned}$$

(where the map is extended linearly to general tensors). This is literally the tensor abstraction of the matrix transpose operation; if  $A = A_i^j w_i \otimes \alpha^j$ , then the dual  $A$  is  $A^* = A_i^j \alpha^i \otimes w_j$ . The matrix of  $A^*$  is precisely the transpose of the matrix of  $A$  with respect to the relevant bases. The map  $*$  itself can be written as a 4-tensor  $*$   $\in V^* \otimes W \otimes W^* \otimes V$ , where  $A^* = * \cdot_{W \otimes V^*} A$ .

There is a notion of the natural pairing of tensor products, which implements composition and evaluation of linear maps, and can be thought of as a natural generalization of scalar multiplication in a field. If  $U$ ,  $V$ , and  $W$  are each finite-dimensional vector spaces, then the bilinear form

$$\begin{aligned} (U \otimes V^*) \times (V \otimes W) &\rightarrow U \otimes \mathbb{R} \otimes W \cong U \otimes W, \\ (u \otimes \alpha, v \otimes w) &\mapsto u \otimes (\alpha \cdot_V v) \otimes w = (\alpha \cdot_V v) u \otimes w \end{aligned}$$

will be denoted also by the infix notation  $\cdot_V$  (i.e.  $(u \otimes \alpha) \cdot_V (v \otimes w) = (\alpha \cdot_V v) u \otimes w$ ). If  $V$  itself is a tensor product of  $n$  factors which are clear from context, then  $\cdot_V$  may be denoted by  $\cdot^n$  (think an  $n$ -fold tensor contraction). If  $n = 2$ , then typically  $:$  is used in place of  $\cdot^2$ . For example, from above,  $A^* = * \cdot_{W \otimes V^*} A = * : A$ .

Given a permutation  $\sigma \in S_n$ , define a right-action by  $\sigma: V_1 \otimes \cdots \otimes V_n \rightarrow V_{\sigma^{-1}(1)} \otimes \cdots \otimes V_{\sigma^{-1}(n)}$ , mapping elements in the obvious way. For example,  $(2\ 3\ 4)$  acting on  $v_1 \otimes v_2 \otimes v_3 \otimes v_4$  puts the second factor in the third position, the third factor in the fourth position, and the fourth factor in the second, giving  $v_1 \otimes v_4 \otimes v_2 \otimes v_3$ . This permutation is itself a linear map and of course can be written as a tensor. However, because it is defined in terms of a right action, the “domain factors” will come on the left. Thus  $\sigma$  is written as a tensor of the form  $V_1^* \otimes \cdots \otimes V_n^* \otimes V_{\sigma^{-1}(1)} \otimes \cdots \otimes V_{\sigma^{-1}(n)}$  (i.e. as a  $2n$ -tensor). Certain tensor constructions are conducive to using such permutations. In the above example,  $*$  can be written as  $(1\ 2) \in W^* \otimes V \otimes V^* \otimes W$ .

The permutation right-action also works naturally when notated using superscripts. For example, if  $B \in U \otimes V \otimes W$ , then

$$B^{(1\ 2)} := B \cdot_{U^* \otimes V^* \otimes W^*} (1\ 2) \in V \otimes U \otimes W$$

and so

$$\begin{aligned} \left( B^{(1\ 2)} \right)^{(2\ 3)} &= (B \cdot_{U^* \otimes V^* \otimes W^*} (1\ 2)) \cdot_{V^* \otimes U^* \otimes W^*} (2\ 3) \\ &= B \cdot_{U^* \otimes V^* \otimes W^*} ((1\ 2) \cdot_{V^* \otimes U^* \otimes W^*} (2\ 3)) \\ &= B \cdot_{U^* \otimes V^* \otimes W^*} (1\ 2) (2\ 3) \\ &= B \cdot_{U^* \otimes V^* \otimes W^*} (1\ 3\ 2) \in V \otimes W \otimes U. \end{aligned}$$

When multiplying the permutations  $(1\ 2)$  and  $(2\ 3)$  in the third line, it is important to note that they are read left-to-right, since they are acting on  $B$  on the right.

The inline cycle notation is somewhat ambiguous in isolation because the number of factors in the domain/codomain is not specified, let alone their types. This information can sometimes be inferred from context, such as from the natural pairing subscripts, as in the following examples.

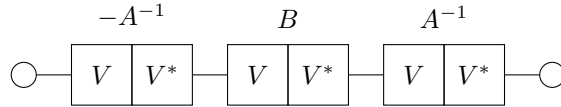
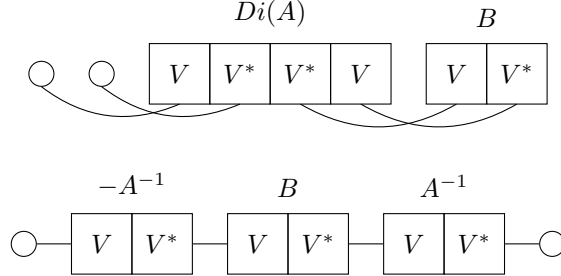
**Example 3.2** (Linearizing the inversion map). Let  $i: GL(V) \rightarrow GL(V)$ ,  $A \mapsto A^{-1}$ , i.e. the linear map inversion operator, where  $GL(V)$  is an open submanifold of  $V \otimes V^*$  via the isomorphism  $V \otimes V^* \cong \text{Hom}(V, V)$ . Its linearization (derivative)  $Di: GL(V) \rightarrow V \otimes V^* \otimes (V \otimes V^*)^* \cong V \otimes V^* \otimes V^* \otimes V$  at  $A \in GL(V)$  in the



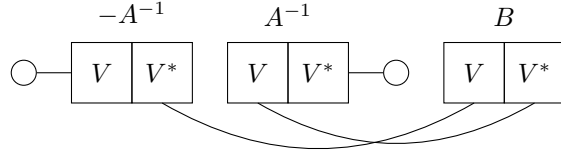
direction  $B \in T_A (GL (V)) \cong V \otimes V^*$  is

$$\begin{aligned}
Di(A) \cdot_{V \otimes V^*} B &= Di \cdot_{V \otimes V^*} \delta(A + \epsilon B) \\
&= \delta(i(A + \epsilon B)) \\
&= \delta\left((A + \epsilon B)^{-1}\right) \\
&= \delta\left(\left((1 + \epsilon BA^{-1}) A\right)^{-1}\right) \\
&= \delta\left(A^{-1} (1 + \epsilon BA^{-1})^{-1}\right) \\
&= \delta\left(A^{-1} \sum_{n=0}^{\infty} (-\epsilon BA^{-1})^n\right) \\
&\quad (|-\epsilon BA^{-1}| \text{ is taken arbitrarily small due to the derivative } \delta := \frac{d}{d\epsilon} |_{\epsilon=0} \\
&\quad \text{being evaluated in an arbitrarily small neighborhood of } \epsilon = 0) \\
&= \delta\left(A^{-1} - \epsilon A^{-1} B A^{-1} + O(\epsilon^2)\right) \\
&= -A^{-1} \cdot_V B \cdot_V A^{-1}.
\end{aligned}$$

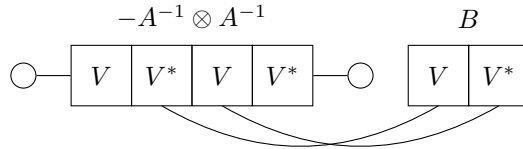
In order to “move” the  $B$  parameter out so that it plays the same syntactical role as in the original expression  $Di(A) \cdot B$ , via adjacent natural pairing, some simple tensor manipulations can be done. The process is easily and accurately expressed via diagram. The following sequence of diagrams is a sequence of equalities. The diagram should be self-explanatory, but for reference, the number of boxes for a particular label denotes the rank of the tensor, with each box labeled with its type. The lines connecting various boxes are natural pairings, and the circles represent the unpaired “slots”, which comprise the type of the resulting expression.



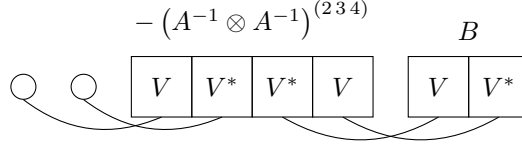
The following step is nothing but moving the boxes for  $B$  out; the natural pairings still apply to the same slots, hence the cables dangling below.



In this setting, a tensor product amounts to flipantly gluing boxes together.



In order for  $B$  to be naturally paired in the same adjacent manner as in the original expression  $Di(A) \cdot B$ , the slots of  $-A^{-1} \otimes A^{-1}$  must be permuted; the second moves to the third, the third to the fourth, and the fourth to the second.



The first diagram equals the last one, thus  $Di(A) \cdot_{V \otimes V^*} B = -(A^{-1} \otimes A^{-1})^{(234)} \cdot_{V \otimes V^*} B$ , and by the nondegeneracy of the natural pairing on  $V \otimes V^*$ , this implies that  $Di(A) = -(A^{-1} \otimes A^{-1})^{(234)}$ , noting that the statement of this expression does not require the direction vector  $B$ . The permutation exponent  $(234)$  can be calculated easily using simple tensors, if not by the above diagrammatic manipulations;

$$(a_1 \otimes a_2) \cdot (b_1 \otimes b_2) \cdot (a_3 \otimes a_4) = (a_1 \otimes a_4 \otimes a_2 \otimes a_3) : (b_1 \otimes b_2) = (a_1 \otimes a_2 \otimes a_3 \otimes a_4)^{(234)} : (b_1 \otimes b_2).$$

Here, the expression  $(a_1 \otimes a_2) \cdot (b_1 \otimes b_2) \cdot (a_3 \otimes a_4)$  represents the expression  $A^{-1} \cdot B \cdot A^{-1}$ .

The next example will later be extended to the setting of Riemannian manifolds and their metric tensor fields, and put to use to formulate what are known as harmonic maps (see (12.7)). But first, a new tensor operation must be defined.

**Definition 3.3** (Parallel tensor product). If  $U, V, W, X$  are vector spaces and  $A \in U \otimes V$  and  $B \in W \otimes X$ , then define their **parallel tensor product**  $A \boxtimes B$  by

$$A \boxtimes B := (A \otimes B)^{(23)} \in (U \otimes W) \otimes (V \otimes X).$$

The parentheses in the type specification are unnecessary, but hint at what the tensor decomposition for the quantity  $A \boxtimes B$  should be, if used as an operand to  $\boxtimes$  again (see below).

If  $A$  and  $B$  represent linear maps, then  $A \boxtimes B \in (U \otimes W) \otimes (V \otimes X)$  represents their tensor product as linear maps (the parentheses are unnecessary but hint at what the domain and codomain are, and for use of  $A \boxtimes B$  as an operand in another parallel tensor product), which is a “parallel” composition; if  $\alpha \in V^*$  and  $\beta \in X^*$ , then  $(A \boxtimes B) \cdot_{V^* \otimes X^*} (\alpha \otimes \beta) = (A \cdot_{V^*} \alpha) \otimes (B \cdot_{X^*} \beta)$ .

There is a slight ambiguity in the notation coming from a lack of specification on how the tensor product of the operands is decomposed in the case when there is more than one such decomposition. Notation explicitly resolving this ambiguity will not be needed in this paper as the relevant tensor product is usually clear from context.

The parallel tensor product is associative; if  $Y$  and  $Z$  are also vector spaces and  $C \in Y \otimes Z$ , then

$$(A \boxtimes B) \boxtimes C = A \boxtimes (B \boxtimes C) \in (U \otimes W \otimes Y) \otimes (V \otimes X \otimes Z),$$

allowing multiply-parallel tensor products.

**Example 3.4** (Tensor product of inner product spaces). If  $(V, g)$  and  $(W, h)$  are inner product spaces (noting that  $g \in V^* \otimes V^*$  and  $h \in W^* \otimes W^*$  are symmetric, i.e. literally invariant under  $(12)$ ), then  $W \otimes V^*$  is an inner product space having induced inner product  $k(A, B) := \text{tr}_V (g^{-1} \cdot_{V^*} A^* \cdot_{W^*} h \cdot_W B)$ . Here, the “inputs” of  $A$  and  $B$  (the  $V^*$  factors) are being paired using  $g^{-1} \in V \otimes V$ , while the “outputs” (the  $W$  factors) are being paired using  $h \in W^* \otimes W^*$ , and the trace is used to “complete the cycle” by plugging the output into the input, thereby producing a real number. The expression  $k(A, B)$  can be written in a more natural way, which takes advantage of the linear composition, as  $A : k : B$  (or, pedantically,  $A \cdot_{W^* \otimes V} k \cdot_{W \otimes V^*} B$ ), instead of the more common but awkward trace expression mentioned earlier. In the tensor formalism, the inner product  $k$  should have type  $W^* \otimes V \otimes W^* \otimes V$ . Permuting the middle two components of the 4-tensor

$h \otimes g^{-1} \in W^* \otimes W^* \otimes V \otimes V$  gives the correct type. In fact,  $k = h \boxtimes g^{-1}$ . A further advantage to this formulation is that if any or all of  $A, k, B$  are functions, there is a clear product rule for derivatives of the expression  $A : k : B$ . This is something that is used critically in Riemannian geometry in the form of covariant derivatives of tensor fields (see (8.2)).

In this paper, the main use of the tensor formulation of linear maps is twofold: to facilitate linear algebraic constructions which would otherwise be difficult or awkward (this includes the ability to express derivatives of [possibly vector or manifold-valued] maps without needing to “plug in” the derivative’s directional argument), and to make clear the product-rule behavior of many important differentiable constructions.

## 4 Bundle Constructions

In order to use the calculus of variations involving Lagrangians depending tangent maps of maps between smooth manifolds, it suffices to consider Lagrangians defined on smooth vector bundle morphisms. Continuing in the style of the previous section, a “full” tensor product of smooth vector bundles (4.4) will be formulated which will then allow expression of smooth vector bundle morphisms as tensor fields, sometimes called two-point tensor fields [11, pg. 70]. The full arsenal of tensor calculus can then be used to considerable advantage.

First, some definitions and simpler bundle constructions will be introduced. A **smooth [fiber] bundle** (hereafter referred to simply as a smooth bundle) is a 4-tuple  $(\mathcal{E}, E, \pi, N)$  where  $\mathcal{E}$ ,  $E$  and  $N$  are smooth manifolds and  $\pi : E \rightarrow N$  is locally trivial, i.e.  $N$  is covered by open sets  $\{U_\alpha\}$  such that  $\pi^{-1}(U_\alpha) \cong U_\alpha \times \mathcal{E}$  as smooth manifolds. The manifolds  $\mathcal{E}$ ,  $E$  and  $N$  are called the **typical fiber**, the **total space**, and the **base space** respectively. The map  $\pi$  is called the **bundle projection**. The full 4-tuple specifying a bundle can be recovered from the bundle projection map, so a locally trivial smooth map can be said to define a smooth bundle. The dimension of the typical fiber of a bundle will be called its **rank**, and will be denoted by  $\text{rank } \pi$  or  $\text{rank } E$  when the bundle is understood from context.

The **space of smooth sections** of a smooth bundle defined by  $\pi : E \rightarrow N$  is

$$\Gamma(\pi) := \{\sigma \in C^\infty(N, E) \mid \pi \circ \sigma = \text{Id}_N\},$$

and may also be denoted by  $\Gamma(E)$ , if the bundle is clear from context. If nonempty,  $\Gamma(\pi)$  is generally an infinite-dimensional manifold (the exception being when the base space  $N$  is finite).

**Proposition 4.1** (Trivial bundle). *Let  $M$  and  $N$  be smooth manifolds. With  $M \rtimes N := M \times N$  and*

$$\pi^{M \rtimes N} := \text{pr}_2^{M \times N} : M \rtimes N \rightarrow N$$

*defines a smooth bundle  $(M, M \rtimes N, \pi^{M \rtimes N}, N)$ , called a **trivial bundle**. Similarly, with  $M \rtimes N := M \times N$  and  $\pi^{M \rtimes N} := \text{pr}_1^{M \times N} : M \rtimes N \rightarrow M$ ,  $(N, M \rtimes N, \pi^{M \rtimes N}, M)$  is a trivial bundle.*

No proof is deemed necessary for (4.1), as each bundle projection trivializes globally in the obvious way. The  $\rtimes$  symbol is a composite of  $\times$  (indicating direct product) and  $\rightarrow$  or  $\leftarrow$  (indicating the base space).

If  $M$  and  $N$  are smooth manifolds as in (4.1), then there are two particularly useful natural identifications.

$$\begin{array}{ll} C^\infty(M, N) \cong \Gamma(M \rtimes N) & C^\infty(M, N) \cong \Gamma(N \rtimes M) \\ \phi \mapsto \text{Id}_M \times_M \phi & \phi \mapsto \phi \times_M \text{Id}_M \\ \text{pr}_2^{M \times N} \circ \Phi \leftarrow \Phi & \text{pr}_1^{N \times M} \circ \Phi \leftarrow \Phi \end{array}$$

These identifications can be thought of identifying a map  $\phi \in C^\infty(M, N)$  with its graph in  $M \times N$  and  $N \times M$  respectively. Furthermore, this allows bundle theory to be applied to reasoning about spaces of maps. The symbols  $M \rtimes N$  and  $N \rtimes M$  now carry a significant amount of meaning. Generally  $N \rtimes M$

will be used in this paper, for consistency with the  $\text{Hom}(V, W) \cong W \otimes V^*$  convention discussed in Section 3. The symbols  $\rtimes$  and  $\ltimes$  are examples of telescoping notation, as they are built notationally on  $\times$ , and conceptually on the direct product, which is what is denoted by  $\times$ . The arrow portion of the symbols can be discarded when type-specificity is not needed.

**Proposition 4.2** (Direct product bundle). *Let  $(\mathcal{E}, E, \pi^E, M)$  and  $(\mathcal{F}, F, \pi^F, N)$  be smooth bundles. Then*

$$\pi^E \times \pi^F : E \times F \rightarrow M \times N, (e, f) \mapsto (\pi^E(e), \pi^F(f))$$

*defines a smooth bundle  $(\mathcal{E} \times \mathcal{F}, E \times F, \pi^E \times \pi^F, M \times N)$ . This bundle is called the **direct product of  $\pi^E$  and  $\pi^F$** , and is not necessarily a trivial bundle.*

*Proof.* Let  $\Psi^E : (\pi^E)^{-1}(U) \rightarrow U \times \mathcal{E}$  and  $\Psi^F : (\pi^F)^{-1}(V) \rightarrow V \times \mathcal{F}$  trivialize  $\pi^E$  and  $\pi^F$  over open sets  $U \subseteq M$  and  $V \subseteq N$  respectively. Then

$$\Psi^E \times \Psi^F : (\pi^E)^{-1}(U) \times (\pi^F)^{-1}(V) \rightarrow U \times \mathcal{E} \times V \times \mathcal{F}$$

has inverse  $(\Psi^E)^{-1} \times (\Psi^F)^{-1}$ . Note that

$$\begin{aligned} (\pi^E)^{-1}(U) \times (\pi^F)^{-1}(V) &= \{(e, f) \in E \times F \mid \pi^E(e) \in U, \pi^F(f) \in V\} \\ &= \{(e, f) \in E \times F \mid (\pi^E \times \pi^F)(e, f) \in U \times V\} \\ &= (\pi^E \times \pi^F)^{-1}(U \times V), \end{aligned}$$

and that

$$P : (U \times \mathcal{E}) \times (V \times \mathcal{F}) \rightarrow (U \times V) \times (\mathcal{E} \times \mathcal{F}), ((u, e), (v, f)) \mapsto ((u, v), (e, f))$$

defines a diffeomorphism. Then

$$\Psi^{E \times F} := P \circ (\Psi^E \times \Psi^F) : (\pi^E \times \pi^F)^{-1}(U \times V) \rightarrow (U \times V) \times (\mathcal{E} \times \mathcal{F})$$

defines a diffeomorphism, and

$$\begin{aligned} & \text{pr}_1^{(U \times V) \times (\mathcal{E} \times \mathcal{F})} \circ \Psi^{E \times F}(e, f) \\ &= \text{pr}_1^{(U \times V) \times (\mathcal{E} \times \mathcal{F})} \circ P \circ (\Psi^E \times \Psi^F)(e, f) \\ &= \text{pr}_1^{(U \times V) \times (\mathcal{E} \times \mathcal{F})} \circ P(\Psi^E(e), \Psi^F(f)) \\ &= \text{pr}_1^{(U \times V) \times (\mathcal{E} \times \mathcal{F})} \circ P(\Psi^E(e), \Psi^F(f)) \\ &= \text{pr}_1^{(U \times V) \times (\mathcal{E} \times \mathcal{F})} \circ P((\text{pr}_1^{U \times \mathcal{E}} \circ \Psi^E(e), \text{pr}_2^{U \times \mathcal{E}} \circ \Psi^E(e)), (\text{pr}_1^{V \times \mathcal{F}} \circ \Psi^F(f), \text{pr}_2^{V \times \mathcal{F}} \circ \Psi^F(f))) \\ &= \text{pr}_1^{(U \times V) \times (\mathcal{E} \times \mathcal{F})}((\text{pr}_1^{U \times \mathcal{E}} \circ \Psi^E(e), \text{pr}_1^{V \times \mathcal{F}} \circ \Psi^F(f)), (\text{pr}_2^{U \times \mathcal{E}} \circ \Psi^E(e), \text{pr}_2^{V \times \mathcal{F}} \circ \Psi^F(f))) \\ &= (\text{pr}_1^{U \times \mathcal{E}} \circ \Psi^E(e), \text{pr}_1^{V \times \mathcal{F}} \circ \Psi^F(f)) \\ &= (\pi^E(e), \pi^F(f)) \\ &= (\pi^E \times \pi^F)(e, f), \end{aligned}$$

showing that  $\Psi^{E \times F}$  trivializes  $\pi^E \times \pi^F$  over  $U \times V \subseteq M \times N$ . Since  $M \times N$  can be covered by such trivializing sets, this establishes that  $\pi^E \times \pi^F$  defines a smooth bundle. The typical fiber of  $\pi^E \times \pi^F$  is  $\mathcal{E} \times \mathcal{F}$ .  $\square$

A **smooth vector bundle** is a fiber bundle whose typical fiber is a vector space and whose local trivializations are linear isomorphisms when restricted to each fiber. If  $(\mathcal{E}, E, \pi, M)$  is a smooth vector bundle, then its **dual vector bundle**  $(\mathcal{E}^*, E^*, \pi^*, M)$  is a smooth vector bundle defined in the following way.

$$E^* := \coprod_{p \in M} (E_p)^*, \quad \pi^* : E^* \rightarrow M, \eta_p \mapsto p.$$

Because  $\mathcal{E}$  is a vector space, the notation  $\mathcal{E}^*$  is already defined. In analogy with Section 3, there are natural pairings on a vector bundle and its dual, defined simply by evaluation. If  $p \in M$ ,  $\eta \in E_p^*$  and  $e \in E_p$ , then  $\eta \cdot_E e := \eta \cdot_{E_p} e$  and  $e \cdot_E \eta := e \cdot_{E_p} \eta$ . Both expressions evaluate to  $\eta(e)$ . Natural traces and  $n$ -fold tensor contraction can be defined analogously. Again, while seemingly pedantic, the subscripted natural pairing notation will prove to be a valuable tool in articulating and error-checking calculations involving vector bundles. To generalize the rest of Section 3 will require the definition of additional structures.

For the remainder of this section, let  $(\mathcal{E}, E, \pi^E, M)$  and  $(\mathcal{F}, F, \pi^F, N)$  now be smooth *vector* bundles. The following construction is essentially an alternate notation for  $\pi^E \times \pi^F: E \times F \rightarrow M \times N$ , but is one that takes advantage of the fact that  $\pi^E$  and  $\pi^F$  are vector bundles, and encodes in the notation the fact that the resulting construction is also a vector bundle. This is analogous to how  $V \times W$  is a vector space with a natural structure if  $V$  and  $W$  are vector spaces, except that this is usually denoted by  $V \oplus W$ .

**Proposition 4.3** (“Full” direct sum vector bundle). *If*

$$E \oplus_{M \times N} F := E \times F,$$

*Then*

$$\pi^E \oplus_{M \times N} \pi^F := \pi^E \times \pi^F: E \oplus_{M \times N} F \rightarrow M \times N$$

*defines a smooth vector bundle  $(\mathcal{E} \oplus \mathcal{F}, E \oplus_{M \times N} F, \pi^E \oplus_{M \times N} \pi^F, M \times N)$ , called the **full direct sum** of  $\pi^E$  and  $\pi^F$ .*

*For each  $(p, q) \in M \times N$ , the vector space structure on  $(\pi^E \oplus_{M \times N} \pi^F)^{-1}(p, q)$  is given in the following way. Let  $\alpha \in \mathbb{R}$  and  $(e_1, f_1), (e_2, f_2) \in (\pi^E \oplus_{M \times N} \pi^F)^{-1}(p, q)$ . Then*

$$\alpha(e_1, f_1) + (e_2, f_2) = (\alpha e_1 + e_2, \alpha f_1 + f_2).$$

*It is critical to see (4.5) for remarks on notation.*

*Proof.* Let  $U, V, \mathcal{E}, \mathcal{F}, P, \Psi^E, \Psi^F$  and  $\Psi^{E \times F}$  be as in the proof of (4.2), and define  $\Psi^{E \oplus_{M \times N} F} := \Psi^{E \times F}$ . Noting that  $\Psi^{E \oplus_{M \times N} F}$  is a smooth bundle isomorphism over  $\text{Id}_{U \times V}$ , so to show that  $\Psi^{E \oplus_{M \times N} F}$  is a linear isomorphism in each fiber, it suffices to show that it is linear in each fiber. Let  $\alpha \in \mathbb{R}$ ,  $(p, q) \in U \times V$  and  $(e_1, f_1), (e_2, f_2) \in (\pi^E \oplus_{M \times N} \pi^F)^{-1}(p, q)$ . Then

$$\begin{aligned} \Psi^{E \oplus_{M \times N} F}(\alpha e_1 + e_2, \alpha f_1 + f_2) &= P \circ (\Psi^E \times \Psi^F)(\alpha e_1 + e_2, \alpha f_1 + f_2) \\ &= P(\Psi^E(\alpha e_1 + e_2), \Psi^F(\alpha f_1 + f_2)) \\ &= P(\alpha \Psi^E(e_1) + \Psi^E(e_2), \alpha \Psi^F(f_1) + \Psi^F(f_2)) \\ &\quad \text{(by trivial vector bundle structures on } U \times \mathcal{E} \text{ and } V \times \mathcal{F}) \\ &= \alpha P(\Psi^E(e_1), \Psi^F(f_1)) + P(\Psi^E(e_2), \Psi^F(f_2)) \\ &\quad \text{(by trivial vector bundle structure on } (U \times V) \times (\mathcal{E} \times \mathcal{F})) \\ &= \alpha P \circ (\Psi^E \times \Psi^F)(e_1, f_1) + P \circ (\Psi^E \times \Psi^F)(e_2, f_2) \\ &= \alpha \Psi^{E \oplus_{M \times N} F}(e_1, f_1) + \Psi^{E \oplus_{M \times N} F}(e_2, f_2). \end{aligned}$$

Thus  $\Psi^{E \oplus_{M \times N} F}$  is linear in each fiber, and because it is invertible, it is a linear isomorphism in each fiber. In particular,  $\Psi^{E \oplus_{M \times N} F}$  is a smooth vector bundle isomorphism over  $\text{Id}_{U \times V}$ . Applying  $(\Psi^{E \oplus_{M \times N} F})^{-1}$  to the above equation gives

$$(\alpha e_1 + e_2, \alpha f_1 + f_2) = \alpha(e_1, f_1) + (e_2, f_2),$$

as desired. □

This construction differs from the Whitney sum of two vector bundles, as the base spaces of the bundles are kept separate, and aren’t even required to be the same. This allows the identification of  $T(M \times N) \rightarrow M \times N$

as  $TM \oplus_{M \times N} TN \rightarrow M \times N$ , which may be done without comment later in this paper. Some important related structures are  $\text{pr}_1^* \pi_M^{TM} : \text{pr}_1^* TM \rightarrow M \times N$  and  $\text{pr}_2^* \pi_N^{TN} : \text{pr}_2^* TN \rightarrow M \times N$ , where  $\text{pr}_i := \text{pr}_i^{M \times N}$ .

The next construction is what will be used in the implementation of smooth vector bundle morphisms as tensor fields.

**Proposition 4.4** (“Full” tensor product bundle). *If*

$$E \otimes_{M \times N} F := \coprod_{(p,q) \in M \times N} E_p \otimes F_q \text{ (disjoint union),}$$

Then

$$\pi^E \otimes_{M \times N} \pi^F : E \otimes_{M \times N} F \rightarrow M \times N, \alpha^{ij} e_i \otimes f_j \mapsto (\pi^E(e_i), \pi^F(f_j)) \text{ (here, } \alpha^{ij} \in \mathbb{R})$$

defines a smooth vector bundle  $(\mathcal{E} \otimes \mathcal{F}, E \otimes_{M \times N} F, \pi^E \otimes_{M \times N} \pi^F, M \times N)$ , called the **full tensor product**<sup>2</sup> of  $\pi^E$  and  $\pi^F$ .

It is critical to see (4.5) for remarks on notation.

*Proof.* Since the argument  $\alpha^{ij} e_i \otimes f_j$  in the definition of  $\pi^E \otimes_{M \times N} \pi^F$  is not necessarily unique, the well-definedness of  $\pi^E \otimes_{M \times N} \pi^F$  must be shown. Let  $\alpha^{ij} e_i^1 \otimes f_j^1 = \beta^{ij} e_i^2 \otimes f_j^2$ . Then in particular,  $\alpha^{ij} e_i^1 \otimes f_j^1, \beta^{ij} e_i^2 \otimes f_j^2 \in E_p \otimes F_q$  for some  $(p, q) \in M \times N$ , and therefore  $e_i^1, e_i^2 \in E_p$  and  $f_j^1, f_j^2 \in F_q$  for each index  $i$  and  $j$ . Thus  $\pi^E(e_i^1) = p = \pi^E(e_i^2)$  and  $\pi^F(f_j^1) = q = \pi^F(f_j^2)$ , so the expression defining  $\pi^E \otimes_{M \times N} \pi^F$  is well-defined.

The set  $E \otimes_{M \times N} F$  does not have an a priori global smooth manifold structure, as it is defined as the disjoint union of vector spaces. A smooth manifold structure compatible with that of the constituent vector spaces will now be defined.

Let  $\Psi^E : (\pi^E)^{-1}(U) \rightarrow U \times \mathcal{E}$  and  $\Psi^F : (\pi^F)^{-1}(V) \rightarrow V \times \mathcal{F}$  trivialize  $\pi^E$  and  $\pi^F$  over open sets  $U \subseteq M$  and  $V \subseteq N$  respectively, such that  $\Psi^E$  and  $\Psi^F$  are each linear in each fiber. Define

$$\begin{aligned} \Psi^{E \otimes_{M \times N} F} : (\pi^E \otimes_{M \times N} \pi^F)^{-1}(U \times V) &\rightarrow (U \times V) \times (\mathcal{E} \otimes \mathcal{F}) \\ X &\mapsto ((\pi^E \otimes_{M \times N} \pi^F)(X), ((\text{pr}_2^{U \times \mathcal{E}} \circ \Psi^E) \otimes (\text{pr}_2^{V \times \mathcal{F}} \circ \Psi^F))(X)). \end{aligned}$$

The map  $\Psi^{E \otimes_{M \times N} F}$  is well-defined and smooth in each fiber by construction, since for each  $(p, q) \in U \times V$ ,

$$(\text{pr}_2^{U \times \mathcal{E}} \circ \Psi^E) \otimes (\text{pr}_2^{V \times \mathcal{F}} \circ \Psi^F) |_{E_p \otimes F_q} : E_p \otimes F_q \rightarrow \mathcal{E} \otimes \mathcal{F}$$

is a linear isomorphism by construction. Additionally,  $\Psi^{E \otimes_{M \times N} F}$  has been constructed so that

$$\text{pr}_1^{(U \times V) \times (\mathcal{E} \otimes_{M \times N} \mathcal{F})} \circ \Psi^{E \otimes_{M \times N} F} = \pi^E \otimes_{M \times N} \pi^F$$

on  $(\pi^E \otimes_{M \times N} \pi^F)^{-1}(U \times V)$ . Define the smooth structure on  $(\pi^E \otimes_{M \times N} \pi^F)^{-1}(U \times V) \subseteq E \otimes_{M \times N} F$  by declaring  $\Psi^{E \otimes_{M \times N} F}$  to be a diffeomorphism. The map  $\pi^E \otimes_{M \times N} \pi^F$  is trivialized over  $U \times V$ . The set  $E \otimes_{M \times N} F$  can be covered by such trivializing open sets. Thus  $E \otimes_{M \times N} F$  has been shown to be locally diffeomorphic to the direct product of smooth manifolds, and therefore it has been shown to be a smooth manifold. With respect to the smooth structure on  $E \otimes_{M \times N} F$ , the map  $\pi^E \otimes_{M \times N} \pi^F$  is smooth, and has therefore been shown to define a smooth vector bundle.  $\square$

*Remark 4.5* (Notation regarding base space). The “full” direct sum (4.3) and “full” tensor product (4.4) bundle constructions allow direct sums and tensor products to be taken of vector bundles when the base spaces differ. If the base spaces are the same, then the construction “joins” them, producing a vector bundle over that shared base space. For example, if  $E$  and  $F$  are vector bundles over  $M$ , then  $E \otimes_{M \times M} F$  has base space  $M \times M$ , while  $E \otimes F$  has base space  $M$ . The base space can be specified in either case as a notational aide; the latter example would be written as  $E \otimes_M F$ . If no subscript is provided on the  $\otimes$  symbol, then

<sup>2</sup>This construction is alluded to in [7, pg. 121], but is not defined or discussed.

the base spaces are “joined” if possible (if they are the same space), otherwise they are kept separate, as in the “full” tensor product construction. This notational convention conforms to the standard Whitney sum and tensor product bundle notation, and uses the notion of telescoping notation to provide more specificity when necessary.

Given a fiber bundle, a natural vector bundle can be constructed “on top” of it, essentially quantifying the variations of bundle elements along each fiber. This is known as the vertical bundle, and it plays a critical role in the development of Ehresmann connections, which provide the “horizontal complement” to the vertical bundle.

**Proposition 4.6** (Vertical bundle). *Let  $\pi^E: E \rightarrow M$  define a smooth [fiber] bundle. If  $VE := \ker T\pi^E \leq TE$ , then  $\pi^{VE} := \pi_E^{TE} \mid_{VE}: VE \rightarrow E$  defines a smooth vector bundle subbundle of  $\pi_E^{TE}: TE \rightarrow E$ , called the **vertical bundle** over  $E$ . Furthermore, the fiber over  $e \in E$  is  $V_e E = T_e E_{\pi^E(e)} \leq T_e E$ .*

*Proof.* Because  $\pi^E$  is a smooth surjective submersion,  $VE \rightarrow E$  is a subbundle of  $TE \rightarrow E$  having corank  $\dim M$  and therefore rank equal to that of  $E$ . Furthermore, if  $e \in E$  and  $\epsilon \mapsto e_\epsilon \in E_{\pi^E(e)}$ , then  $\delta e_\epsilon$  represents an arbitrary element of  $T_e E_{\pi^E(e)}$ , and  $T\pi^E(\delta e_\epsilon) = \delta(\pi^E(e_\epsilon)) = \delta(\pi(e)) = 0$ , showing that  $\delta e_\epsilon \in \ker T\pi^E$ , and therefore that  $\delta e_\epsilon \in V_e E$ . This shows that  $T_e E_{\pi^E(e)} \subseteq V_e E$ . Because  $\dim T_e E_{\pi^E(e)} = \text{rank } E$ , this shows that  $T_e E_{\pi^E(e)} = V_e E$ .  $\square$

Given the extra structure that a vector bundle provides over a [fiber] bundle, there is a canonical smooth vector bundle isomorphism which adds significant value to the pullback bundle formalism used throughout this paper. This can be seen put to greatest use in Part II, for example, in development of the first variation (see (12.1)).

**Proposition 4.7** (Vertical bundle as pullback). *If  $\pi: E \rightarrow M$  defines a smooth vector bundle, then*

$$\begin{aligned} \iota_{VE}^{\pi^* E}: \pi^* E &\rightarrow VE, \\ (x, y) &\mapsto \delta_\epsilon(x + \epsilon y) \end{aligned}$$

*is a smooth vector bundle isomorphism over  $\text{Id}_E$ , called the **vertical lift**, having inverse*

$$\iota_{\pi^* E}^{VE}: \delta e_\epsilon \mapsto \left( e_0, \lim_{\epsilon \rightarrow 0} \frac{e_\epsilon - e_0}{\epsilon} \right),$$

*where, without loss of generality,  $e_\epsilon$  is an  $E$ -valued variation which lies entirely in a single fiber.*

*Proof.* It is clear that  $\iota_{VE}^{\pi^* E}$  is linear and injective on each fiber. By a dimension counting argument, it is therefore an isomorphism on each fiber. Because it preserves the basepoint, it is a vector bundle isomorphism over  $\text{Id}_E$ . Because the map  $(x, y, \epsilon) \mapsto x + \epsilon y$  is smooth, so is the defining expression for  $\iota_{VE}^{\pi^* E}$ , thereby establishing smoothness. That  $\iota_{\pi^* E}^{VE}$  inverts  $\iota_{VE}^{\pi^* E}$  is a trivial calculation.  $\square$

## 5 Strongly-Typed Tensor Field Operations

Because vector bundles and the related operations can be thought of conceptually as “sheaves of linear algebra”, the constructions in Section 3, generalized earlier in this section, can be further generalized to the setting of sections of vector bundles.

If  $E, F, G$  are smooth vector bundles over  $M$ , then define the natural pairing of a tensor field with a vector:

$$\begin{aligned} \cdot_F: \Gamma(E \otimes_M F^*) \times F &\rightarrow E, \\ (e \otimes_M \phi, f) &\mapsto e(\pi^F(f)) [\phi(\pi^F(f)) \cdot_F f], \end{aligned}$$

extending linearly to general tensor fields. Further, define the natural pairing of tensor fields:

$$\begin{aligned} \cdot_F : \Gamma(E \otimes_M F^*) \times \Gamma(F \otimes_M G) &\rightarrow \Gamma(E \otimes_M G), \\ (e \otimes_M \phi, f \otimes_M g) &\mapsto (p \mapsto e(p) \otimes_M (\phi(p) \cdot_{F_p} f(p)) \otimes_M g(p)) \\ &= (p \mapsto (\phi(p) \cdot_{F_p} f(p)) (e \otimes_M g)(p)), \end{aligned}$$

extending linearly to general tensor fields. This multiple use of the  $\cdot_F$  symbol is a concept known as **operator overloading** in computer programming. No ambiguity is caused by this overloading, as the particular use can be inferred from the types of the operands. As before, the subscript  $F$  may be optionally omitted when clear from context.

The permutations defined in Section 3 are generalized as tensor fields. If  $F_1, \dots, F_n$  are smooth vector bundles over  $M$ , and  $\sigma \in S_n$  is a permutation, then  $\sigma$  can act on  $F_1 \otimes_M \dots \otimes_M F_n$  by permuting its factors, and therefore can be identified with a tensor field

$$\sigma \in \Gamma(F_1^* \otimes_M \dots \otimes_M F_n^* \otimes_M F_{\sigma^{-1}(1)} \otimes_M \dots \otimes_M F_{\sigma^{-1}(n)})$$

defined by

$$(f_1 \otimes_M \dots \otimes_M f_n) \cdot_{F_1^* \otimes_M \dots \otimes_M F_n^*} \sigma := f_{\sigma^{-1}(1)} \otimes_M \dots \otimes_M f_{\sigma^{-1}(n)}.$$

An important feature of such permutation tensor fields is that they are parallel with respect to covariant derivatives on the factors  $F_1, \dots, F_n$  (see (8.12) for more on this).

## 6 Pullback Bundles

The pullback bundle, defined below, is a crucial building block for many important bundle constructions, as it enriches the type system dramatically, and allows the tensor formulation of linear algebra to be extended to the vector bundle setting. In particular, the abstract, global formulation of the space of smooth vector bundle morphisms over a map  $\phi: M \rightarrow N$  is achieved quite cleanly using a pullback bundle. Furthermore, the use of pullback bundles and pullback covariant derivatives simplifies what would otherwise be local coordinate calculations, thereby giving more insight into the geometric structure of the problem.

For the duration of this section, let  $(\mathcal{F}, F, \pi, N)$  be a smooth bundle having rank  $r$ .

**Proposition 6.1** (Pullback bundle). *Let  $M$  and  $N$  be smooth manifolds and let  $\phi: M \rightarrow N$  be smooth. If*

$$\phi^* F := \{(m, f) \in M \times F \mid \phi(m) = \pi(f)\},$$

and

$$\pi^{\phi^* F} := \text{pr}_1^{M \times F} |_{\phi^* F} : \phi^* F \rightarrow M, (m, f) \mapsto m,$$

then  $(\mathcal{F}, \phi^* F, \pi^{\phi^* F}, M)$  defines a smooth bundle. In particular,  $\phi^* F$  is a smooth manifold having dimension  $\dim M + \text{rank } \pi$ . The bundle defined by  $\pi^{\phi^* F}$  is called the **pullback of  $\pi$  by  $\phi$** .

*Proof.* Recalling that  $\mathcal{F}$  denotes the typical fiber of  $\pi$ , let  $\Psi: \pi^{-1}(U) \rightarrow U \times \mathcal{F}$  trivialize  $\pi$  over open set  $U \subseteq N$ . Define

$$\Psi_\phi: \phi^* (\pi^{-1}(U)) \rightarrow \phi^{-1}(U) \times \mathcal{F}, (m, f) \mapsto (m, \text{pr}_2^{U \times \mathcal{F}} \circ \Psi(f))$$

and

$$\Psi_\phi^{-1}: \phi^{-1}(U) \times \mathcal{F} \rightarrow \phi^* (\pi^{-1}(U)), (m, f) \mapsto (m, \Psi^{-1}(\phi(m), f)).$$

Claim (1):  $\Psi_\phi$  and  $\Psi_\phi^{-1}$  are smooth. Proof:  $\phi^* (\pi^{-1}(U)) \subseteq \phi^{-1}(U) \times \pi^{-1}(U)$ , and  $\Psi_\phi$  is clearly smooth as a map defined on the larger manifold. Therefore it restricts to a smooth map on  $\phi^* (\pi^{-1}(U))$ . An analogous argument shows that  $\Psi_\phi^{-1}$  is smooth. Claim (1) proved.



Claim (2):  $\Psi_\phi^{-1}$  inverts  $\Psi_\phi$ . Proof: Let  $(m, f) \in \phi^*(\pi^{-1}(U))$ . Then

$$\begin{aligned}\Psi_\phi^{-1} \circ \Psi_\phi(m, f) &= \Psi_\phi^{-1}(m, \text{pr}_2^{U \times \mathcal{F}} \circ \Psi(f)) \\ &= (m, \Psi^{-1}(\phi(m), \text{pr}_2^{U \times \mathcal{F}} \circ \Psi(f))) \\ &= (m, \Psi^{-1}(\pi(f), \text{pr}_2^{U \times \mathcal{F}} \circ \Psi(f))) \quad (\text{since } \phi(m) = \pi(f)) \\ &= (m, \Psi^{-1}(\text{pr}_1^{U \times \mathcal{F}} \circ \Psi(f), \text{pr}_2^{U \times \mathcal{F}} \circ \Psi(f))) \\ &= (m, \Psi^{-1} \circ \Psi(f)) \\ &= (m, f).\end{aligned}$$

With  $g \in \mathcal{F}$ ,

$$\begin{aligned}\Psi_\phi \circ \Psi_\phi^{-1}(m, g) &= \Psi_\phi(m, \Psi^{-1}(\phi(m), g)) \\ &= (m, \text{pr}_2^{U \times \mathcal{F}} \circ \Psi \circ \Psi^{-1}(\phi(m), g)) \\ &= (m, \text{pr}_2^{U \times \mathcal{F}}(\phi(m), g)) \\ &= (m, g),\end{aligned}$$

proving Claim (2).

Claim (3):  $\Psi_\phi$  trivializes  $\pi^{\phi^*F}$  over  $\phi^{-1}(U) \subseteq M$ . Proof: Let  $(m, f) \in \phi^*(\pi^{-1}(U))$ . Then

$$\text{pr}_1^{\phi^{-1}(U) \times \mathcal{F}} \circ \Psi_\phi(m, f) = \text{pr}_1^{\phi^{-1}(U) \times \mathcal{F}} \circ (m, \text{pr}_2^{U \times \mathcal{F}} \circ \Psi(f)) = m = \pi^{\phi^*F}(m, f),$$

and by claims (1) and (2),  $\Psi_\phi$  is a diffeomorphism, so  $\Psi_\phi$  trivializes  $\pi^{\phi^*F}$  over  $\phi^{-1}(U) \subseteq M$ . Claim (3) proved.

Since  $M$  can be covered with sets as in claim (3) and since the typical fiber of  $\pi^{\phi^*F}$  is diffeomorphic to  $\mathcal{F}$ , this shows that  $\pi^{\phi^*F}$  defines a smooth bundle  $(\mathcal{F}, \phi^*F, \pi^{\phi^*F}, M)$ . Because  $\phi^*F$  is locally diffeomorphic to the product of an open subset of  $M$  with  $\mathcal{F}$ ,  $\phi^*F$  has been shown to be a smooth manifold having dimension  $\dim M + \dim \mathcal{F} = \dim M + \text{rank } \pi$ .  $\square$

While the pullback bundle is constructed as a submanifold of a direct product, there is a natural bundle morphism into the pulled-back bundle, which serves as an interface to maps defined on the pulled-back bundle. Usually this morphism is notationally suppressed, just as naturally isomorphic spaces can be identified without explicit notation.

**Corollary 6.2** (Pullback fiber projection bundle morphism). *If  $\phi: M \rightarrow N$  is smooth, then*

$$\begin{aligned}\rho_F^{\phi^*F}: \phi^*F &\rightarrow F, \\ (m, f) &\mapsto f\end{aligned}$$

*is a smooth bundle morphism over  $\phi$  which is an isomorphism when restricted to any fiber of  $\phi^*F$ .*

Because  $\rho_F^{\phi^*F}$  is the projection  $\text{pr}_F^{M \times F}|_{\phi^*F}$ , its tangent map is also just the projection  $\text{pr}_{TF}^{TM \oplus TF}|_{T\phi^*F}$ .

**Proposition 6.3** (Bundle pullback is a contravariant functor). *The map of categories*

$$\begin{aligned}\text{Pullback: Manifold} &\rightarrow \{\text{Bundle}(M) \mid M \in \text{Manifold}\}, \\ M &\mapsto \text{Bundle}(M), \\ (\phi: M \rightarrow N) &\mapsto \left(\text{Bundle}(N) \rightarrow \text{Bundle}(M), (\mathcal{F}, F, \pi, N) \mapsto (\mathcal{F}, \phi^*F, \pi^{\phi^*F}, M)\right)\end{aligned}$$

*is a contravariant functor. Here, naturally isomorphic bundles in  $\text{Bundle}(M)$ , for each manifold  $M$ , are identified (along with the corresponding morphisms).*

*Proof.* Noting that

$$\text{Id}_N^* F = \{(n, f) \in N \times F \mid \text{Id}_N(n) = \pi(f)\} \cong F$$

and that

$$\begin{aligned} (\text{Id}_N^* \pi)(n, f) &= (\text{pr}_1^{N \times F} \mid_{\text{Id}_N^* F})(n, f) = n = \pi(f) \\ \implies \text{Id}_N^* \pi &\cong \pi, \end{aligned}$$

it follows that  $\text{Pullback}(\text{Id}_N) = \text{Id}_{\text{Bundle}(N)} = \text{Id}_{\text{Pullback}(N)}$ , i.e. Pullback satisfies the identity axiom of functoriality.

For the contravariance axiom, let  $\phi: M \rightarrow N$  and  $\psi: L \rightarrow M$  be smooth manifold morphisms and let  $(\mathcal{F}, F, \pi, N)$  be a smooth bundle. Then

$$\begin{aligned} \psi^* \phi^* F &= \{(\ell, p) \in L \times \phi^* F \mid \psi(\ell) = \pi^{\phi^* F}(p)\} \\ &= \{(\ell, (m, f)) \in L \times (M \times F) \mid \psi(\ell) = \pi^{\phi^* F}(m, f) \text{ and } \phi(m) = \pi(f)\} \\ &= \{(\ell, (m, f)) \in L \times (M \times F) \mid \psi(\ell) = m \text{ and } \phi(m) = \pi(f)\} \\ &\cong \{(\ell, f) \in L \times F \mid \phi \circ \psi(\ell) = \pi(f)\} \\ &= (\phi \circ \psi)^* F \end{aligned}$$

and

$$\begin{aligned} \pi^{\psi^* \phi^* F}(\ell, (m, f)) &= (\text{pr}_1^{L \times \phi^* F} \mid_{\psi^* \phi^* F})(\ell, (m, f)) = \ell \text{ and} \\ \pi^{(\phi \circ \psi)^* F}(\ell, f) &= (\text{pr}_1^{L \times F} \mid_{(\phi \circ \psi)^* F})(\ell, f) = \ell, \end{aligned}$$

showing that  $\pi^{\psi^* \phi^* F} \cong \pi^{(\phi \circ \psi)^* F}$ , and therefore

$$\text{Pullback}(\psi) \circ \text{Pullback}(\phi) = \text{Pullback}(\phi \circ \psi),$$

establishing Pullback as a contravariant functor.  $\square$

The space of sections of a pullback bundle is easily quantified.

$$\Gamma(\phi^* F) = \left\{ \sigma \in C^\infty(M, \phi^* F) \mid \pi^{\phi^* F} \circ \sigma = \text{Id}_M \right\}.$$

This space will be central in the theory developed in the rest of this paper. Furthermore, it is naturally identified with the space of sections along the pullback map;

$$\Gamma_\phi(F) := \left\{ \Sigma \in C^\infty(M, F) \mid \pi^F \circ \Sigma = \phi \right\}.$$

These spaces are naturally isomorphic to one another, and therefore an identification can be made when convenient. While the former space is more correct from a strongly typed standpoint, the latter space is a convenient and intuitive representational form. The particular correspondence depends heavily on the fact that  $\phi^* F$  is a submanifold of  $M \times F$ .

$$\begin{aligned} \Gamma(\phi^* F) &\cong \Gamma_\phi(F) \\ \sigma &\mapsto \text{pr}_2^{M \times F} \circ \sigma, \\ \text{Id}_M \times_M \Sigma &\leftarrow \Sigma. \end{aligned}$$

Furthermore, if  $f \in \Gamma(F)$ , then  $f \circ \phi \in \Gamma_\phi(F)$ . Note that it is *not* true that any  $\sigma \in \Gamma_\phi(F)$  can be written as  $f \circ \phi$  for some  $f \in \Gamma(F)$ , for example when there exists some distinct  $p, q \in M$  such that  $\phi(p) = \phi(q)$  and  $\sigma(p) \neq \sigma(q)$ . Furthermore, the representation  $f \circ \phi$  is generally non-unique, for example when  $\phi$  is not surjective, sections  $f_1, f_2 \in \Gamma(F)$  which differ only away from the image of  $\phi$  will still give  $f_1 \circ \phi = f_2 \circ \phi$ . Before developing the notion of a linear connection on a pullback bundle, it will be necessary to address these features which, while inconvenient, provide the strength of the pullback bundle and pullback covariant derivative (see (8.8)).

**Lemma 6.4** (Local representation of  $\Gamma_\phi(F)$  elements). *Recall that  $r$  denotes the rank of smooth bundle  $F$ . If  $\sigma \in \Gamma_\phi(F)$  then each point  $p \in M$  has some neighborhood  $U$  in which  $\sigma$  can be written locally as  $\sigma|_U = \sigma^i f_i \circ \phi|_U$ , where  $f_1, \dots, f_r \in \Gamma(F|_{\phi(U)})$  is a frame for  $F|_{\phi(U)}$ , and  $\sigma^1, \dots, \sigma^r \in C^\infty(U, \mathbb{R})$  are defined by  $\sigma^i = (f^i \circ \phi|_U) \cdot_F \sigma|_U$ .*

*Proof.* Let  $p \in M$ , let  $V \subseteq N$  be a neighborhood of  $\phi(p)$  over which  $F|_V$  is trivial, and let  $U = \phi^{-1}(V)$ , so that  $U$  is a neighborhood of  $p$ . Let  $f_1, \dots, f_r \in \Gamma(F|_V)$  be a frame for  $F|_V$  (i.e.  $F|_{\phi(U)}$ ), and let  $f^1, \dots, f^r \in \Gamma((F|_V)^*)$  be the corresponding coframe (i.e. the unique  $f^1, \dots, f^r$  such that  $f^i \cdot_F f_j = \delta_j^i$  for each  $i, j$ ). Define  $\sigma^i \in C^\infty(M, \mathbb{R})$  by  $\sigma^i = (f^i \circ \phi|_U) \cdot_F \sigma|_U$ . Then

$$\begin{aligned} \sigma^i f_i \circ \phi|_U &= (f^i \circ \phi|_U) \cdot_F \sigma|_U f_i \circ \phi|_U \\ &= ((f_i \circ \phi|_U) \otimes_U (f^i \circ \phi|_U)) \cdot_F \sigma|_U \\ &= ((f_i \otimes_V f^i) \circ \phi|_U) \cdot_F \sigma|_U \\ &= (\text{Id}_{F|_V} \circ \phi|_U) \cdot_F \sigma|_U \\ &= \sigma|_U, \end{aligned}$$

as desired.  $\square$

Some literature uses expressions of the form  $f \circ \phi \in \Gamma_\phi(F)$  along with an implicit use of the section-identifying isomorphism to write down particular sections of pullback bundles. In most cases, this tacit identification of spaces is harmless, but certain highly involved calculations may suffer from it. The section that  $f \circ \phi$  corresponds to under said isomorphism is  $\text{Id}_M \times_M (f \circ \phi) \in \Gamma(\phi^*F)$ . However, because this expression is unwieldy and therefore a more compact and contextually meaningful expression is called for.

**Definition 6.5** (Pullback section). If  $f \in \Gamma(F)$  and  $\phi: M \rightarrow N$  is smooth, then define

$$\phi^*f := \text{Id}_M \times_M (f \circ \phi) \in \Gamma(\phi^*F).$$

This is known as a **pullback section**.

The pullback section is deservedly named. If  $\phi: M \rightarrow N$  and  $\psi: L \rightarrow M$  are smooth, then  $\psi^*\phi^*f \cong (\phi \circ \psi)^*f$  in the sense of the proof of (6.3).

**Proposition 6.6** (Bundle pullback commutes with tensor product). *If  $E$  and  $F$  are smooth vector bundles over manifold  $N$  and  $\phi: M \rightarrow N$  is smooth, then the map*

$$\begin{aligned} \phi^*E \otimes_M \phi^*F &\rightarrow \phi^*(E \otimes_N F), \\ (m, e) \otimes_M (m, f) &\mapsto (m, e \otimes_N f) \end{aligned}$$

(extended linearly to general tensors) is a smooth vector bundle isomorphism.

*Proof.* Let  $c$  denote the above map. The well-definedness of  $c$  comes from the universal mapping property on multilinear forms which induces a linear map on a corresponding tensor product. If  $c((m, e) \otimes_M (m, f)) = 0$ , then  $e \otimes_N f = 0$ , which implies that  $e = 0$  or  $f = 0$ , and therefore that  $(m, e) \otimes_M (m, f) = 0$ . Because there exists a basis for  $(\phi^*E \otimes_M \phi^*F)_m$  consisting only of simple tensors, this implies that  $c$  is injective, and by a dimensionality argument, that  $c$  is an isomorphism. The map is clearly smooth and respects the fiber structures of its domain and codomain. Thus  $c$  is a smooth vector bundle isomorphism.  $\square$

The contravariance of pullback and its naturality with respect to tensor product are two essential properties which provide some of the flexibility and precision of the strongly typed tensor formalism described in this paper. This will become quite apparent in Part II.

*Remark 6.7* (Tensor field formulation of smooth vector bundle morphisms). A particularly useful application of pullback bundles is in forming a rich type system for smooth vector bundle morphisms. This approach was inspired by [20, pg. 11]. Let  $\pi^E: E \rightarrow M$  and  $\pi^F: F \rightarrow N$  be smooth vector bundles, and let  $\phi: M \rightarrow N$  be smooth. Consider  $\text{Hom}_\phi(E, F)$ , i.e. the space of smooth vector bundle morphisms over the map  $\phi$ . There is a natural identification with another space which lets the base map  $\phi$  play a more direct role in the space's type. In particular,

$$\begin{aligned} \text{Hom}_\phi(E, F) &\cong \text{Hom}_{\text{Id}_M}(E, \phi^*F), \\ A &\mapsto \pi^E \times_E A, \\ \text{pr}_2^{M \times F} \circ B &\leftarrow B. \end{aligned}$$

This particular identification of smooth vector bundle morphisms over  $\phi$  can now be directly translated into the tensor field formalism, analogously to (3.1).

$$\begin{aligned} \Gamma(\phi^*F \otimes_M E^*) &\rightarrow \text{Hom}_{\text{Id}_M}(E, \phi^*F), \\ A &\mapsto (e \mapsto A \cdot_E e). \end{aligned}$$

The inverse image of  $B \in \text{Hom}_{\text{Id}_M}(E, \phi^*F)$  is given locally; let  $(e_i)$  and  $(f_i)$  denote local frames for  $E$  and  $F$  in neighborhoods  $U \subseteq M$  and  $V \subseteq N$  respectively, with  $\phi(U) \subseteq V$ , and let  $(e^i)$  and  $(f^i)$  denote their dual coframes. Then the tensor field corresponding to  $B$  is given locally in  $U$  by  $B_j^i \phi^* f_i \otimes_M e^j$ , where  $B_j^i := \phi^* df^i \circ B \circ e_j \in C^\infty(U, \mathbb{R})$ .

Quantifying smooth vector bundle morphisms as the tensor fields lends itself naturally to doing calculus on vector and tensor bundles, as the relevant derivatives (covariant derivatives) take the form of tensor fields. The type information for a particular vector bundle morphism is encoded in the relevant tensor bundle.

## 7 Tangent Map as a Tensor Field

This section deals specifically with the tangent map operator by using concepts from Section 5 and Section 6 to place it in a strongly typed setting and to prepare to unify a few seemingly disparate concepts and notation for some tangible benefit (in particular, see Section 10).

Given a smooth map  $\phi: M \rightarrow N$ , its tangent map  $T\phi: TM \rightarrow TN$  is a smooth vector bundle morphism over  $\phi$ , so by (6.7), is naturally identified with a tensor field

$$\nabla^{M \rightarrow N} \phi \in \Gamma(\phi^*TN \otimes_M T^*M),$$

which may be denoted by  $\nabla \phi$  where type pedantry is deemed unnecessary. This construction is known as a **two-point tensor field** [11, pg. 70]. The inscribed  $\circ$  symbol in  $\nabla$  is used to denote that this is a nonlinear derivative, thereby distinguishing it from a linear covariant derivative.

*Remark 7.1* (Generalized covariant derivative). The well-known one-to-one correspondence between linear connections and linear covariant derivatives [8, pg. 520] generalizes to a one-to-one correspondence between Ehresmann connections and a generalized notion of covariant derivative. To give a partial definition for the purposes of utility, a **generalized covariant derivative** on a smooth [fiber] bundle  $F \rightarrow N$  is a map  $\nabla$  on  $\Gamma(F)$  such that  $\nabla \sigma \in \Gamma(\sigma^*VF \otimes_N T^*N)$  for each  $\sigma \in \Gamma(F)$ . The space of maps  $C^\infty(M, N)$  is naturally identified as  $\Gamma(N \rtimes M)$ , and there is a natural Ehresmann connection on the bundle  $N \rtimes M$ , whose corresponding covariant derivative is the tangent map operator. This is the subject of another of the author's papers and will not be discussed here further. This is mentioned here to incorporate linear covariant derivatives (to be introduced and discussed in Section 8) and the tangent map operator (a nonlinear covariant derivative) under the single category “covariant derivative”.

There is a subtle issue regarding construction of the cotangent map of  $\phi$  which is handled easily by the tensor field construction. In particular, while the cotangent map  $T^*\phi$  is the pointwise adjoint of the tangent map  $T\phi$ , i.e. for each  $p \in M$ ,  $T_p\phi: T_pM \rightarrow T_{\phi(p)}N$  is linear and  $T_p^*\phi: T_{\phi(p)}^*N \rightarrow T_p^*M$  is the adjoint of  $T_p\phi$ , it does not follow that  $T^*\phi \in \text{Hom}(T^*N, T^*M)$ , being some sort of “total adjoint” of  $T\phi \in \text{Hom}(TM, TN)$ . The obstruction is due to the fact that  $\phi$  may not be surjective, so there may be some fiber  $T_q^*N$  that is not of the form  $T_{\phi(p)}^*N$ , and therefore the domain could not be all of  $T^*N$ . Furthermore, even if  $\phi$  were surjective, if it weren’t also injective, say  $\phi(p_0) = \phi(p_1)$  for some distinct  $p_0, p_1 \in M$ , then  $T_{\phi(p_0)}^*N = T_{\phi(p_1)}^*N$ , and  $T_{p_0}M \neq T_{p_1}M$ , so the action on the fiber  $T_{\phi(p_0)}^*N$  is not well-defined.

In the tensor field parlance, the cotangent map  $T^*\phi$  simply takes the form

$$(\nabla\phi)^{(12)} \in \Gamma(T^*M \otimes_M \phi^*TN).$$

The permutation superscript (12) is used here instead of  $*$  to distinguish it notationally from pullback notation, which will be necessary in later calculations. The key concept is that the tensor field  $(\nabla\phi)^{(12)}$  encodes the base map  $\phi$ ; the basepoint  $p \in M$  is part of the domain  $\phi^*T^*N$  itself.

The chain rule in the tensor field formalism makes use of the bundle pullback. If  $\psi: L \rightarrow M$  is smooth, then

$$\nabla^{L \rightarrow N}(\phi \circ \psi) = \psi^* \nabla^{M \rightarrow N} \phi \cdot \psi^* T_M \nabla^{L \rightarrow M} \psi.$$

Because  $\nabla\psi \in \Gamma(\psi^*TM \otimes_L T^*L)$ , to form a well-defined natural pairing, the use of the pullback

$$\psi^* \nabla\phi \in \Gamma(\psi^*(\phi^*TN \otimes_M T^*M)) = \Gamma(\psi^*\phi^*TN \otimes_L \psi^*T^*M) = \Gamma((\phi \circ \psi)^*TN \otimes_L \psi^*T^*M)$$

is necessary (instead of just  $\nabla\phi \in \Gamma(\phi^*TN \otimes_M T^*M)$ ).

Sometimes it is useful to discard some type information and write  $\nabla\phi \in \Gamma_{\phi \times_M \text{Id}_M}(TN \otimes_{N \times M} T^*M)$ , i.e.  $\nabla\phi: M \rightarrow TN \otimes_{N \times M} T^*M$  such that  $(\pi_N^{TN} \otimes_{N \times M} \pi_M^{T^*M}) \circ \nabla\phi = \phi \times_M \text{Id}_M$ . This is easily done by the canonical fiber projection available to all pullback bundle constructions;  $\phi^*TN \otimes_M T^*M \cong (\phi \times_M \text{Id}_M)^*(TN \otimes_{N \times M} T^*M)$ , and the canonical fiber projection is

$$\rho_{TN \otimes_{N \times M} T^*M}^{(\phi \times_M \text{Id}_M)^*(TN \otimes_{N \times M} T^*M)}: (\phi \times_M \text{Id}_M)^*(TN \otimes_{N \times M} T^*M) \rightarrow TN \otimes_{N \times M} T^*M,$$

as defined in (6.2). The granularity of the type system should reflect the weight of the calculations being performed. For demonstration of contrasting situations, see the discussion at the beginning of Section 8 and the computation of the first variation in (12.1).

It is important to have notation which makes the distinction between the smooth vector bundle morphism formalism and the tensor field formalism, because it may sometimes be necessary to mix the two, though this paper will not need this. An added benefit to the tensor field formulation of tangent maps is that certain notions regarding derivatives can be conceptually and notationally combined, for example in Section 10.

## 8 Linear Covariant Derivatives

As will be shown in the following discussion, a linear covariant derivative (commonly referred to in the standard literature without the “linear” qualifier) provides a way to generalize the notion in elementary calculus of the differential of a vector-valued function. The linear covariant derivative interacts naturally with the notion of the pullback bundle, and this interaction leads naturally to what could be called a covariant derivative chain rule, which provides a crucial tool for the tensor calculus computations seen later.

Let  $V$  and  $W$  be finite-dimensional vector spaces let  $U \subseteq V$  be open, and let  $\phi: U \rightarrow W$  be differentiable. Recall from elementary calculus the differential  $D\phi: U \rightarrow W \otimes V^*$  (essentially matrix-valued). There

is no base map information encoded in  $D\phi$  (i.e.  $\phi$  can't be recovered from  $D\phi$  alone), it contains only derivative information. The vector space structure of  $V$  and  $W$  allows the trivializations  $TU \cong V \rtimes U$  and  $TW \cong W \rtimes W$ , where the first factors are the base spaces and the second factors are the fibers (see (4.1)). The tangent map  $\nabla^{U \rightarrow W} \phi: U \rightarrow TW \otimes_{W \times U} T^*U$  (see Section 7) has a codomain that can be trivialized similarly;

$$TW \otimes_{W \times U} T^*U \cong (W \rtimes W) \otimes_{W \times U} (V^* \rtimes U) \cong (W \otimes V^*) \rtimes (W \times U).$$

Because  $(W \otimes V^*) \rtimes (W \times U)$ , as a set, is a direct product, it can be decomposed into two factors. Letting  $\text{pr}_1$  and  $\text{pr}_2$  be the projections onto the first and second factors respectively,

$$\text{pr}_1 \circ \nabla \phi: U \rightarrow W \otimes V^* \text{ and } \text{pr}_2 \circ \nabla \phi: U \rightarrow W \times U.$$

The map  $\text{pr}_2 \circ \nabla \phi$  is the element of  $\Gamma(W \rtimes U)$  identified with the base map  $\phi$  itself;  $\text{pr}_1^{W \times U} \circ \text{pr}_2 \circ \nabla \phi = \phi$ . This base map information is discarded in defining the differential of  $\phi$  as  $D\phi := \text{pr}_1 \circ \nabla \phi$ ; the fiber portion of  $\nabla \phi$ . This construction relies critically on the natural isomorphism  $TW \cong W \rtimes W$  for a vector space  $W$ .

An analogous construction shows that the differential  $D\phi$  of a map  $\phi$  is well-defined even when its domain is a manifold. However, when the codomain of a map  $\phi$  is only a manifold, there does not in general exist a natural trivialization of its tangent bundle (in contrast to the vector space case), and therefore  $D\phi$  can't be defined without additional structure. A linear covariant derivative provides the missing structure.

For the remainder of this section, let  $\pi: E \rightarrow N$  define a smooth vector bundle having rank  $r$ .

A linear covariant derivative on  $E$  provides a means of taking derivatives of sections of  $E$  (i.e. maps  $\sigma: N \rightarrow E$  such that  $\pi \circ \sigma = \text{Id}_N$ ) without passing to a higher tangent bundle as would happen under the tangent map functor (i.e. if  $\sigma \in \Gamma(E)$  then  $T\sigma: TN \rightarrow TE$  and  $\nabla^{N \rightarrow E} \sigma: N \rightarrow TE \otimes_{E \times N} T^*N$ ). A linear covariant derivative provides an effective “trivialization” of  $TE$  analogous to the trivialization  $TW \cong W \rtimes W$  as discussed above, discarding all but the “fiber” portion of the derivative, allowing the construction of an object known as the total linear covariant derivative analogous to the differential  $D\phi$  as discussed above.

The notion of a linear covariant derivative on a vector bundle is arguably the crucial element of differential geometry<sup>3</sup>. In particular, this operator implements the product rule property common to anything that can be called a derivation – a property which is particularly conducive to the operation of tensor calculus. The total linear covariant derivative of a vector field (i.e. section of a vector bundle) allows the generalization of many constructions in elementary calculus to the setting of smooth vector bundles equipped with linear covariant derivatives. For example, the divergence  $\text{div } X := \text{tr } DX$  of a vector field  $X$  on  $\mathbb{R}^n$  generalizes to the divergence  $\text{div } X := \text{tr } \nabla X$  of a vector field  $X$  on  $N$ , which has an analogous divergence theorem among other qualitative similarities.

*Remark 8.1* (Natural linear covariant derivative on trivial line bundle). Before making the general definition for the linear covariant derivative, a natural linear covariant derivative will be introduced. With  $N$  denoting a smooth manifold as before, if  $f \in C^\infty(N, \mathbb{R})$ , then  $df \in \Gamma(T^*N)$  is the **differential** of  $f$ . Let

$$\nabla^{N \rightarrow \mathbb{R}} f := df.$$

Because  $C^\infty(N, \mathbb{R})$  is naturally identified with  $\Gamma(\mathbb{R} \rtimes N)$ , this is essentially the natural linear covariant derivative on the trivial line bundle  $\mathbb{R} \rtimes N$ . Note that there is an associated product rule; if  $f, g \in C^\infty(N, \mathbb{R})$ , then  $fg \in C^\infty(N, \mathbb{R})$ , and

$$\nabla^{N \rightarrow \mathbb{R}}(fg) = d(fg) = gdf + f dg = g \nabla^{N \rightarrow \mathbb{R}} f + f \nabla^{N \rightarrow \mathbb{R}} g.$$

When clear from context, the superscript decoration can be omitted and the derivative denoted as  $\nabla f$ .

<sup>3</sup>The *Fundamental Lemma of Riemannian Geometry* establishes the existence of the Levi-Civita connection[9, pg. 68], which is a linear covariant derivative satisfying certain naturality properties.

**Definition 8.2** (Linear covariant derivative). A **linear covariant derivative** on a vector bundle defined by  $\pi: E \rightarrow N$  is an  $\mathbb{R}$ -linear map  $\nabla^E: \Gamma(E) \rightarrow \Gamma(E \otimes_N T^*N)$  satisfying the product rule

$$\nabla^E(f \otimes_N \sigma) = \sigma \otimes_N \nabla^{N \rightarrow \mathbb{R}} f + f \otimes_N \nabla^E \sigma, \quad (8.1)$$

where  $f \in C^\infty(N, \mathbb{R})$  and  $\sigma \in \Gamma(E)$ . The switch in order in the first term of the expression is necessary to form a tensor field of the correct type,  $\Gamma(E \otimes_N T^*N)$ . If  $\sigma \in \Gamma(E)$ , then the expression  $\nabla^E \sigma$  is known as the **total [linear] covariant derivative** of  $\sigma$ . If  $\nabla^E \sigma = 0$  [in a subset  $U \subseteq N$ ], then  $\sigma$  is said to be **parallel** [on  $U$ ]. The “linear” qualifier is implied in standard literature and is therefore often omitted.

The inscribed  $\iota$  in  $\nabla$  is to indicate that the covariant derivative is linear, and can be omitted when clear from context, or when it is unnecessary to distinguish it from the nonlinear tangent map operator whose decorated symbol is  $\nabla^\flat$ . For the remainder of this section, this distinction will not be necessary, so an undecorated  $\nabla$  will be used.

For  $V \in \Gamma(TN)$ , it is customary to denote  $\nabla^E \sigma \cdot V$  by  $\nabla_V^E \sigma$ , where  $V$  indicates the “directional” component of the derivative. Following this convention, the product rule can be written in a form where the product rule is more obvious;

$$\nabla_V^E(f \otimes_N \sigma) = \nabla_V^{N \rightarrow \mathbb{R}} f \otimes_N \sigma + f \otimes_N \nabla_V^E \sigma.$$

A covariant derivative is a local operator with respect to the base space  $N$ ; if  $p \in N$ , then  $(\nabla^E \sigma)(p)$  depends only on the restriction of  $\sigma$  to an arbitrarily small neighborhood of  $p$  [9, pg. 50], and therefore the restriction  $\nabla^E|_U: \Gamma_U(E) \rightarrow \Gamma_U(E \otimes_N T^*N)$  makes sense, allowing calculations using local expressions. Furthermore, a covariant derivative can be constructed locally and glued together under certain conditions. See [8, pg. 503] for more on this, and as a reference for general theory on bundles, covariant derivatives, and connections.

Linear covariant derivatives on several vector bundle constructions will now be developed. In analogy to defining a linear map by its action on a generating subset (e.g. a basis or a dense subspace) and then extending using the linear structure, Lemma (8.6) allows a covariant derivative to be defined on a generating subset (which can be chosen to make the defining expression particularly natural) and then extending. In this case, the relevant space is the space of sections of the vector bundle, which is a module over the ring of smooth functions on a manifold, and the extension process is done via linearity and the product rule (see (8.2)). This approach will allow the local trivialization implementation details to be hidden within the proof of Lemma (8.6) – an example of information hiding – so that constructions of covariant derivatives can proceed clearly by focusing only on the natural properties of the relevant objects and then invoking the lemma to do the “dirty” work (see (8.7) and (8.9)).

A bit of useful notation will be introduced to simplify the next definition. If  $G \subseteq \Gamma$  is a subset of a  $C^\infty(N, \mathbb{R})$ -module  $\Gamma$  whose elements are functions on  $N$  (and therefore have a notion of restriction to a subset) and  $U \subseteq N$  is open, then let  $G_U$  denote the set of restrictions of the elements of  $G$  to the set  $U$ . Note that  $G_U \subseteq \Gamma_U$  by construction.

**Definition 8.3** (Finitely generating subset). Say that a subset of a module **finitely generates** the module if the subset contains a finite set of generators for the module.

**Definition 8.4** (Locally finitely generating subset). If  $\Gamma$  is a  $C^\infty(N, \mathbb{R})$ -module and  $G \subseteq \Gamma$ , then  $G$  is said to be a **locally finitely generating subset** of  $\Gamma$  if each point  $q \in N$  has a neighborhood  $U \subseteq N$  for which  $G_U$  finitely generates  $\Gamma_U$ .

The space of sections of a vector bundle is the archetype for the above definition. The locally trivial nature of  $\pi: E \rightarrow N$  allows local frames to be chosen in a neighborhood of each point of  $N$ , from which global smooth sections (though not necessarily a global frame) can be made using a partition of unity subordinate to the trivializing neighborhoods. The set of such global sections forms a locally finite generating subset of  $\Gamma(E)$ .



**Lemma 8.5.** *If  $G$  is a locally finitely generating subset of  $\Gamma(E)$ , then each point in  $N$  has a neighborhood  $U \subseteq N$  and  $e_1, \dots, e_r \in G_U$  such that  $e_1, \dots, e_r$  forms a frame for  $\Gamma_U(E)$ . In other words, a local frame can be chosen out of  $G$  near each point in  $N$ .*

*Proof.* Let  $q \in N$  and let  $V \subseteq N$  be a neighborhood of  $q$  for which  $G_V = \{g_1, \dots, g_\ell\}$  finitely generates  $\Gamma_V(E)$  (here,  $\ell \geq r$ , recalling that  $r = \text{rank } E$ ). Without loss of generality, let  $g_1(q), \dots, g_r(q)$  be linearly independent (this is possible because  $\{g_1(q), \dots, g_\ell(q)\}$  spans the vector space  $E_q$ ). Because  $g_i$  is continuous for each  $i$  and the linear independence of the sections  $g_1, \dots, g_r$  is an open condition (defined by  $L^{-1}(\mathbb{R} \setminus \{0\})$  where  $L: N \rightarrow \bigwedge^r E_q, p \mapsto g_1(p) \wedge \dots \wedge g_r(p)$ ), there is a neighborhood  $U \subseteq V$  of  $q$  for which  $\{g_1(p), \dots, g_r(p)\}$  is a linearly independent set for each  $p \in U$ . Finally, letting  $e_i := g_i|_U$  for  $i \in \{1, \dots, r\}$ , the sections  $e_1, \dots, e_r \in G_U$  form a frame for  $\Gamma_U(E)$ .  $\square$

The following lemma shows that defining a covariant derivative on a locally finitely generating subset of the space of sections of a vector bundle is sufficient to uniquely define a covariant derivative on the whole space. The particular generating subset can be chosen so the covariant derivative has a particularly natural expression within that subset.

**Lemma 8.6** (Linear covariant derivative construction). *Let  $G$  be a locally finite generating subset of  $\Gamma(E)$ . If  $\nabla^G: G \rightarrow \Gamma(E \otimes_N T^*N)$  satisfies the linear covariant derivative axioms<sup>4</sup>, then there is a unique linear covariant derivative  $\nabla^E: \Gamma(E) \rightarrow \Gamma(E \otimes_N T^*N)$  whose restriction to  $G$  is  $\nabla^G$ .*

*Proof.* If  $q \in N$ , then by (8.5) there exists a neighborhood  $U \subseteq N$  of  $q$  for which there are  $e_1, \dots, e_r \in G_U$  forming a frame for  $E|_U$ . If  $\sigma \in \Gamma(E)$ , then  $\sigma|_U = \sigma^i e_i$  for some  $\sigma^1, \dots, \sigma^r \in C^\infty(U, \mathbb{R})$  (specifically,  $\sigma^i = e^i \cdot_E \sigma|_U$ , where  $e^1, \dots, e^r \in \Gamma_U(E^*)$  denotes the dual coframe of  $e_1, \dots, e_r$ ). Define  $\nabla^E: \Gamma(E) \rightarrow \Gamma(E \otimes_N T^*N)$  locally on  $\Gamma_U(E)$  so as to satisfy the product rule

$$\nabla^E(\sigma|_U) := e_i \otimes_N \nabla^{N \rightarrow \mathbb{R}} \sigma^i + \sigma^i \otimes_N \nabla^G e_i.$$

To show well-definedness, let  $f_1, \dots, f_r \in G_U$  be another frame for  $E|_U$ . Then  $\sigma = \tau^i f_i$  for some  $\tau^1, \dots, \tau^r \in C^\infty(U, \mathbb{R})$ . Let  $\Psi: \Gamma_U(E) \rightarrow \Gamma_U(E)$  be the unique smooth vector bundle isomorphism such that  $f_i = \Psi \cdot_E e_i$ . Writing  $\Psi$  and  $\Psi^{-1}$  with respect to the frame  $(e_i)$  as  $\Psi_j^i e_i \otimes e^j$  and  $(\Psi^{-1})_j^i e_i \otimes e^j$  respectively, it follows that  $f_i = \Psi_j^i e_j$  and  $\tau^i = \sigma^j (\Psi^{-1})_j^i$ . Then

$$\begin{aligned} \nabla^E(\tau^i f_i) &= f_i \otimes_N \nabla^{N \rightarrow \mathbb{R}} \tau^i + \tau^i \otimes_N \nabla^G f_i \\ &= \Psi_j^i e_j \otimes_N \nabla^{N \rightarrow \mathbb{R}} \left( \sigma^k (\Psi^{-1})_k^i \right) + \sigma^j (\Psi^{-1})_j^i \otimes_N \nabla^G (\Psi_i^k e_k) \\ &= \Psi_j^i e_j (\Psi^{-1})_k^i \otimes_N \nabla^{N \rightarrow \mathbb{R}} \sigma^k + \Psi_j^i e_j \sigma^k \otimes_N \nabla^{N \times \mathbb{R}} (\Psi^{-1})_k^i \\ &\quad + \sigma^j (\Psi^{-1})_j^i e_k \otimes_N \nabla^{N \rightarrow \mathbb{R}} \Psi_i^k + \sigma^j (\Psi^{-1})_j^i \Psi_i^k \otimes_N \nabla^G e_k \\ &= \delta_k^j e_j \otimes_N \nabla^{N \rightarrow \mathbb{R}} \sigma^k + \sigma^j \delta_j^k \otimes_N \nabla^G e_k + \sigma^\ell e_k \otimes_N \nabla^{N \rightarrow \mathbb{R}} \left( \Psi_i^k (\Psi^{-1})_\ell^i \right) \\ &= \nabla^E(\sigma^i e_i). \end{aligned}$$

The last equality follows because  $\Psi_i^k (\Psi^{-1})_\ell^i = \delta_\ell^k$ , which is a constant function, so  $\nabla^{N \rightarrow \mathbb{R}} \left( \Psi_i^k (\Psi^{-1})_\ell^i \right) = 0$ . Thus the expression defining  $\nabla^E$  doesn't depend on the choice of local frame. This establishes the well-definedness of  $\nabla^E$ .

Clearly the restriction of  $\nabla^E$  to  $G$  is  $\nabla^G$ . This establishes the claim of existence. Uniqueness follows from the fact that  $\nabla^E$  is defined in terms of the maps  $\nabla^{N \rightarrow \mathbb{R}}$  and  $\nabla^G$ .  $\square$

Lemma (8.6) is used in the proof of the following proposition to allow a natural formulation of the pullback covariant derivative with respect to a natural locally finite generating subset of  $\Gamma(\phi^*E)$ , in which the relevant derivative has a natural chain rule.

<sup>4</sup>What is meant by this is that the product rule must only be satisfied on  $\lambda \otimes_N g$  if  $\lambda g \in G$ , where  $\lambda \in C^\infty(N, \mathbb{R})$  and  $g \in G$ .



**Proposition 8.7** (Pullback covariant derivative). *If  $\phi: M \rightarrow N$  is smooth and  $\nabla^E$  is a covariant derivative on  $E$ , then there is a unique covariant derivative  $\nabla^{\phi^*E}$  on  $\phi^*E$  satisfying the chain rule*

$$\nabla^{\phi^*E} \phi^*e = \phi^* \nabla^E e \cdot_{\phi^*TN} \nabla^{M \rightarrow N} \phi$$

for all  $e \in \Gamma(E)$ .

*Proof.* Let  $G := \{\sigma \in \Gamma(\phi^*E) \mid \sigma = \phi^*e \text{ for some } e \in \Gamma(E)\}$ , noting that a local frame  $e_1, \dots, e_{\text{rank } E} \in \Gamma_U(E)$  over open set  $U \subseteq N$  induces a local frame  $\phi^*e_1, \dots, \phi^*e_{\text{rank } E} \in \Gamma_{\phi^{-1}(U)}(\phi^*E)$ , so  $G$  is a locally finite generating subset of  $\Gamma(\phi^*E)$ . Define

$$\begin{aligned} \nabla^G: G &\rightarrow \Gamma(\phi^*E \otimes_M T^*N), \\ \phi^*e &\mapsto \phi^* \nabla^E e \cdot_{\phi^*TN} \nabla^{M \rightarrow N} \phi. \end{aligned}$$

The well-definedness and  $\mathbb{R}$ -linearity of  $\nabla^G$  comes from that of  $\nabla^E$ . For the product rule, if  $\lambda \in C^\infty(M, \mathbb{R})$  and  $e \in \Gamma(E)$ , then the product  $\lambda \otimes_M \phi^*e$  is an element of  $G$  if and only if  $\lambda = \phi^*\mu$  for some  $\mu \in C^\infty(N, \mathbb{R})$ , in which case,  $\lambda \otimes_M \phi^*e = \phi^*\mu \otimes_M \phi^*e = \phi^*(\mu \otimes_N e)$ . Then it follows that

$$\begin{aligned} \nabla^G(\lambda \otimes_M \phi^*e) &= \nabla^G \phi^*(\mu \otimes_N e) \\ &= \phi^* \nabla^E(\mu \otimes_N e) \cdot_{\phi^*TN} \nabla^{M \rightarrow N} \phi \\ &= \phi^*(e \otimes_N \nabla^{N \rightarrow \mathbb{R}} \mu + \mu \otimes_N \nabla^E e) \cdot_{\phi^*TN} \nabla^{M \rightarrow N} \phi \\ &= \phi^*(e \otimes_N \nabla^{N \rightarrow \mathbb{R}} \mu) \cdot_{\phi^*TN} \nabla^{M \rightarrow N} \phi + \phi^*(\mu \otimes_N \nabla^E e) \cdot_{\phi^*TN} \nabla^{M \rightarrow N} \phi \\ &= \phi^*e \otimes_M (\phi^* \nabla^{N \rightarrow \mathbb{R}} \mu \cdot_{\phi^*TN} \nabla^{M \rightarrow N} \phi) + \phi^*\mu \otimes_M (\phi^* \nabla^E e \cdot_{\phi^*TN} \nabla^{M \rightarrow N} \phi) \\ &= \phi^*e \otimes_M \nabla^{M \rightarrow \mathbb{R}} \phi^*\mu + \phi^*\mu \otimes_M \nabla^G \phi^*e \\ &= \phi^*e \otimes_M \nabla^{M \rightarrow \mathbb{R}} \lambda + \lambda \otimes_M \nabla^G \phi^*e, \end{aligned}$$

which is exactly the required product rule. By (8.6), there exists a unique covariant derivative  $\nabla^{\phi^*E}$  on  $\phi^*E$  whose restriction to  $G$  is  $\nabla^G$ .  $\square$

The full notation  $\nabla^{\phi^*E}$  is often cumbersome, so it may be denoted by  $\nabla^\phi$  when the pulled-back bundle is clear from context.

*Remark 8.8.* There is an important feature of a pullback covariant derivative in the case that pullback map is not an immersion; the pullback covariant derivative may be nonzero even where the pullback map is singular. This fact can be obscured by a certain abuse of notation which often comes in the expression of the geodesic equations in differential geometry (see (12.8)). An example will illustrate this point.

Let  $\nabla^{TM}$  be a covariant derivative on  $\pi_M^{TM}: TM \rightarrow M$ . Let  $\Theta: \mathbb{R} \rightarrow TM$  be a unit-length vector field which describes the location of a person (the basepoint) and direction s/he is looking (the fiber portion) with respect to time (let  $\mathbb{R}$  have standard coordinate  $t$ ). Define  $\theta: \mathbb{R} \rightarrow M$  by  $\theta := \pi_M^{TM} \circ \Theta$ , so that  $\theta$  is the base map of  $\Theta$ , i.e.  $\theta$  has discarded the direction information and only encodes the location information. Say that for some closed interval  $I \subseteq \mathbb{R}$ ,  $\frac{d\theta}{dt}|_I$  is identically zero (and so is not an immersion), but that  $\frac{d\Theta}{dt}|_I$  is nonvanishing; see Figure 8.1. Mathematically, this means that during this time,  $\Theta$  is varying only within a single fiber of  $TM$ . Physically, this means that during this time, the person is standing still but the direction s/he is looking is changing. Passing to a higher tangent space is often undesirable (note that  $\frac{d\Theta}{dt}$  takes values in  $TTM$ ), so to avoid this, a covariant derivative is used. In order to be meaningful, the covariant derivative must capture this fiber-only variation.

Because  $\Theta$  is a vector field along  $\theta$ , it can be written as  $\Theta \in \Gamma(\theta^*TM)$ , and the covariant derivative on  $TM$  induces a pullback covariant derivative on  $\theta^*TM$ , which has base space  $\mathbb{R}$ . In other words,  $\theta^*TM$  is parameterized by time. Then  $\nabla_{\frac{d\theta}{dt}}^{\theta^*TM} \Theta \in \Gamma(\theta^*TM)$  is the desired covariant derivative of  $\Theta$  with respect to time. A coordinate-based calculation will be made to make completely obvious why this pullback covariant derivative captures the desired information. Let  $(x^i)$  be local coordinates on  $M$  and, for simplicity, assume

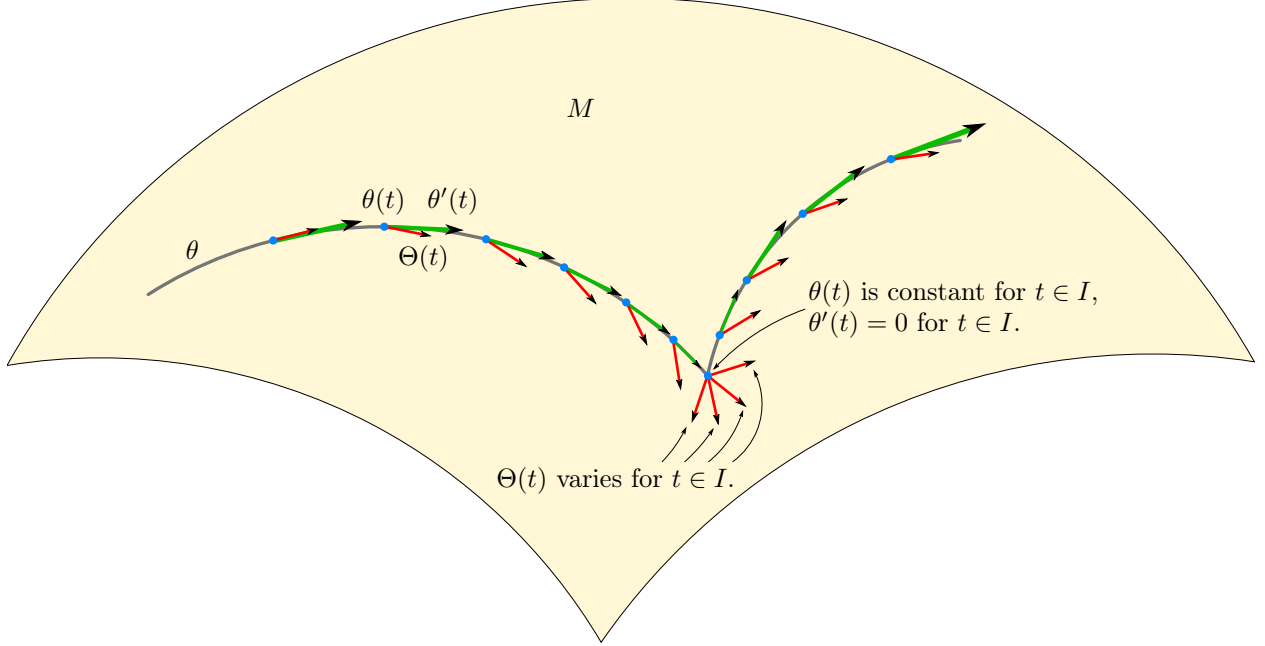


Figure 8.1: A picture of the manifold  $M$ , path  $\theta$ , and vector fields  $\theta'$  and  $\Theta$ . The blue dots represent  $\theta(t)$  at certain points  $t \in \mathbb{R}$ , while the green and red arrows represent  $\theta'(t)$  and  $\Theta(t)$  at these points respectively. Note that  $\Theta$  is a unit-length vector field along  $\theta$  and varies within  $I$ , whereas  $\theta'$  is a vector field along  $\theta$  that vanishes within  $I$ .

that the image of  $\theta$  lies entirely within this coordinate chart. Because  $(\partial_i)$  is a local frame for  $TM$ ,  $(\theta^*\partial_i)$  is a local frame for  $\theta^*TM$ , by (6.4) and  $\Theta \in \Gamma(\theta^*TM)$  can be written locally as  $\Theta(t) = \Theta^i(t) (\theta^*\partial_i)(t)$  for some functions  $(\Theta^i: \mathbb{R} \rightarrow \mathbb{R})$ . Then

$$\begin{aligned}
 \nabla_{\frac{d}{dt}}^{\theta^*TM} \Theta &= \nabla_{\frac{d}{dt}}^{\theta^*TM} (\Theta^i \theta^*\partial_i) \\
 &= \left( \nabla_{\frac{d}{dt}} \Theta^i \right) \theta^*\partial_i + \Theta^i \nabla_{\frac{d}{dt}}^{\theta^*TM} \theta^*\partial_i \\
 &= \frac{d\Theta^i}{dt} \theta^*\partial_i + \Theta^i \theta^* \nabla^{TM} \partial_i \cdot \theta^*TM \nabla^{\mathbb{R} \rightarrow M} \theta \cdot_{T\mathbb{R}} \frac{d}{dt} \\
 &= \frac{d\Theta^i}{dt} \theta^*\partial_i + \Theta^i \theta^* \nabla^{TM} \partial_i \cdot \theta^*TM \frac{d\theta}{dt}.
 \end{aligned}$$

Note that  $\nabla^{\mathbb{R} \rightarrow M} \theta \in \Gamma(\theta^*TM)$ . Within the interval  $I$ ,  $\frac{d\theta}{dt}$  vanishes, so the second term vanishes on  $I$ . However, because  $\Theta$  is varying in a fiber-only direction within  $I$ , the basepoint is not changing and  $\frac{d\Theta^i}{dt} \theta^*\partial_i$  can be identified with an elementary vector space derivative (the fiber is a vector space and so an elementary derivative is well-defined there). This fiber-direction derivative is nonvanishing by assumption, so  $\nabla_{\frac{d}{dt}}^{\theta^*TM} \Theta$  is nonvanishing on  $I$  as desired.

Introducing a bit of natural notation which will be helpful for the next result, if  $X \in \Gamma(E)$  and  $Y \in \Gamma(F)$ , then define  $X \oplus Y \equiv X \oplus_{M \times N} Y \in \Gamma(E \oplus_{M \times N} F)$  and  $X \otimes Y \equiv X \otimes_{M \times N} Y \in \Gamma(E \otimes_{M \times N} F)$  by

$$(X \oplus_{M \times N} Y)(p, q) := X(p) \oplus Y(q) \quad \text{and} \quad (X \otimes_{M \times N} Y)(p, q) := X(p) \otimes Y(q)$$

for each  $(p, q) \in M \times N$ .

**Proposition 8.9** (Induced covariant derivatives on  $E \oplus_{M \times N} F$  and  $E \otimes_{M \times N} F$ ). *If  $\nabla^E$  and  $\nabla^F$  are covariant derivatives on  $E$  and  $F$  respectively, then there are unique covariant derivatives*

$$\nabla^{E \oplus_{M \times N} F}: \Gamma(E \oplus_{M \times N} F) \rightarrow \Gamma((E \oplus_{M \times N} F) \otimes_{M \times N} (T^*M \oplus_{M \times N} T^*N))$$

and

$$\nabla^{E \otimes_{M \times N} F}: \Gamma(E \otimes_{M \times N} F) \rightarrow \Gamma((E \otimes_{M \times N} F) \otimes_{M \times N} (T^*M \oplus_{M \times N} T^*N))$$

on  $E \oplus F$  and  $E \otimes F$  respectively, satisfying the sum rule

$$\nabla_{u \oplus v}^{E \oplus F} (X \oplus Y) = \nabla_u^E X \oplus \nabla_v^F Y$$

and the product rule

$$\nabla_{u \oplus v}^{E \otimes F} (X \otimes Y) = \nabla_u^E X \otimes Y + X \otimes \nabla_v^F Y,$$

respectively, where  $X \in \Gamma(E)$ ,  $Y \in \Gamma(F)$ , and  $u \oplus v \in TM \oplus TN$ . Here,  $TM \oplus TN \rightarrow M \times N$  (and its dual) is used instead of the isomorphic vector bundle  $T(M \times N) \rightarrow M \times N$  (and its dual).

*Proof.* Suppressing the pedantic use of the  $M \times N$  subscript to avoid unnecessary notational overload, the set  $G := \{e \oplus f \mid e \in \Gamma(E), f \in \Gamma(F)\}$  is a locally finite generator of  $\Gamma(E \oplus F)$ , since local frames for  $E \oplus F$  take the form  $\{e_i \oplus 0, 0 \oplus f_j\}$ , where  $\{e_i\}$  and  $\{f_j\}$  are local frames for  $E$  and  $F$  respectively. Define

$$\begin{aligned} \nabla^G: G &\rightarrow \Gamma((E \oplus F) \otimes (T^*M \oplus T^*N)), \\ X \oplus Y &\mapsto (u \oplus v \mapsto \nabla_u^E X \oplus \nabla_v^F Y), \text{ where } u \oplus v \in TM \oplus TN. \end{aligned}$$

This map is well-defined and  $\mathbb{R}$ -linear by construction, since the connections  $\nabla^E$  and  $\nabla^F$  are well-defined and  $\mathbb{R}$ -linear. If  $\lambda \in C^\infty(M \times N, \mathbb{R})$ ,  $X \in \Gamma(E)$ , and  $Y \in \Gamma(F)$ , then the product  $\lambda \otimes (X \oplus Y)$  is in  $G$  (i.e. has the form  $\bar{X} \oplus \bar{Y}$  for some  $\bar{X} \in \Gamma(E)$  and  $\bar{Y} \in \Gamma(F)$ ) if and only if  $\lambda$  is constant. Thus the product rule (restricted to elements of  $G$ ) reduces to  $\mathbb{R}$ -linearity, which is already satisfied. By (8.6), there exists a unique connection  $\nabla^{E \oplus F}$  on  $E \oplus F$  whose restriction to  $G$  is  $\nabla^G$ .

Similarly, the set  $H := \{e \otimes f \mid e \in \Gamma(E), f \in \Gamma(F)\}$  is a locally finite generator of  $\Gamma(E \otimes F)$ , since local frames for  $E \otimes F$  take the form  $\{e_i \otimes f_j\}$ , where  $\{e_i\}$  and  $\{f_j\}$  are local frames for  $E$  and  $F$  respectively. Define

$$\begin{aligned} \nabla^H: H &\rightarrow \Gamma((E \otimes F) \otimes (T^*M \oplus T^*N)), \\ X \otimes Y &\mapsto (u \oplus v \mapsto \nabla_u^E X \otimes Y + X \otimes \nabla_v^F Y), \text{ where } u \oplus v \in TM \oplus TN. \end{aligned}$$

This map is well-defined and  $\mathbb{R}$ -linear by construction, since the connections  $\nabla^E$  and  $\nabla^F$  are well-defined and  $\mathbb{R}$ -linear. For the product rule, with  $\lambda \in C^\infty(M \times N, \mathbb{R})$ ,  $X \in \Gamma(E)$ , and  $Y \in \Gamma(F)$ , the product  $\lambda \otimes (X \otimes Y)$  is in  $H$  if and only if there exist  $\mu \in C^\infty(M, \mathbb{R})$  and  $\nu \in C^\infty(N, \mathbb{R})$  such that  $\lambda = \mu \otimes \nu \in (\mathbb{R} \rtimes M) \otimes (\mathbb{R} \rtimes N)$  (noting that then  $\lambda \otimes_{M \times N} (X \otimes Y) = (\mu \otimes \nu) \otimes_{M \times N} (X \otimes Y) = (\mu \otimes_M X) \otimes (\nu \otimes_N Y)$ ). In this case, with  $u \oplus v \in TM \oplus TN$ ,

$$\begin{aligned} &\nabla_{u \oplus v}^H (\lambda \otimes_{M \times N} (X \otimes Y)) \\ &= \nabla_{u \oplus v}^H ((\mu \otimes \nu) \otimes_{M \times N} (X \otimes Y)) \\ &= \nabla_{u \oplus v}^H ((\mu \otimes_M X) \otimes (\nu \otimes_N Y)) \\ &= \nabla_u^E (\mu \otimes_M X) \otimes (\nu \otimes_N Y) + (\mu \otimes_M X) \otimes \nabla_v^F (\nu \otimes_N Y) \\ &= (\nabla_u^{M \rightarrow \mathbb{R}} \mu \otimes_M X) \otimes (\nu \otimes_N Y) + (\mu \otimes_M \nabla_u^E X) \otimes (\nu \otimes_N Y) \\ &\quad + (\mu \otimes_M X) \otimes (\nabla_v^{N \rightarrow \mathbb{R}} \nu \otimes_N Y) + (\mu \otimes_M X) \otimes (\nu \otimes_N \nabla_v^F Y) \\ &= (\nabla_u^{M \rightarrow \mathbb{R}} \mu \otimes \nu + \mu \otimes \nabla_v^{N \rightarrow \mathbb{R}} \nu) \otimes_{M \times N} (X \otimes Y) + \lambda \otimes_{M \times N} (\nabla_u^E X \otimes Y + X \otimes \nabla_v^F Y) \\ &= \nabla_{u \oplus v}^{M \times N \rightarrow \mathbb{R}} \lambda \otimes_{M \times N} (X \otimes Y) + \lambda \otimes_{M \times N} \nabla_{u \oplus v}^H (X \otimes Y), \end{aligned}$$

which is exactly the required product rule. By (8.6), there exists a unique connection  $\nabla^{E \otimes F}$  on  $E \otimes F$  whose restriction to  $H$  is  $\nabla^H$ .  $\square$

*Remark 8.10* (Naturality of the covariant derivatives on  $E \oplus_{M \times N} F$  and  $E \otimes_{M \times N} F$ ). Letting  $\text{pr}_i := \text{pr}_i^{M \times N}$  ( $i \in \{1, 2\}$ ) for brevity, the maps

$$\begin{aligned} \xi: E \oplus_{M \times N} F &\rightarrow \text{pr}_1^* E \oplus_{M \times N} \text{pr}_2^* F, \\ e \oplus f &\mapsto ((\pi^E \oplus \pi^F)(e \oplus f), e) \oplus_{M \times N} ((\pi^E \oplus \pi^F)(e \oplus f), f) \end{aligned}$$

and

$$\begin{aligned} \psi: E \otimes_{M \times N} F &\rightarrow \text{pr}_1^* E \otimes_{M \times N} \text{pr}_2^* F, \\ e \otimes f &\mapsto ((\pi^E \otimes \pi^F)(e \otimes f), e) \otimes_{M \times N} ((\pi^E \otimes \pi^F)(e \otimes f), f), \end{aligned}$$

each extended linearly to the rest of their domains, are easily shown to be smooth vector bundle isomorphisms over  $\text{Id}_{M \times N}$ . Then

$$\nabla_z^{E \oplus F}(X \oplus Y) = \xi^{-1} \left( \nabla_z^{\text{pr}_1^* E \oplus_{M \times N} \text{pr}_2^* F} \xi(X \oplus Y) \right)$$

and

$$\nabla_z^{E \otimes F}(X \otimes Y) = \psi^{-1} \left( \nabla_z^{\text{pr}_1^* E \otimes_{M \times N} \text{pr}_2^* F} \psi(X \otimes Y) \right)$$

for all  $X \in \Gamma(E)$ ,  $Y \in \Gamma(F)$ , and  $z \in T(M \times N)$ , showing that the connections on  $E \oplus F$  and  $E \otimes F$  are  $\xi$  and  $\psi$ -related to the naturally induced connections on  $\text{pr}_1^* E \oplus \text{pr}_2^* F$  and  $\text{pr}_1^* E \otimes_{M \times N} \text{pr}_2^* F$  respectively, and are therefore in this sense natural. The sum  $X \oplus Y \in \Gamma(E \oplus F)$  and product  $X \otimes Y \in \Gamma(E \otimes F)$  correspond to  $\text{pr}_1^* X \oplus_{M \times N} \text{pr}_2^* Y$  and  $\text{pr}_1^* X \otimes_{M \times N} \text{pr}_2^* Y \in \Gamma(\text{pr}_1^* E \otimes_{M \times N} \text{pr}_2^* F)$  under  $\xi$  and  $\psi$  respectively.

Many important tensor constructions involve permutations. An extremely useful property of these permutations is that they commute with the covariant derivatives induced by the covariant derivatives on the tensor bundle factors, making them natural operators in the setting of covariant tensor calculus.

**Proposition 8.11** (Transposition tensor fields are parallel). *Let  $E_1, E_2, E_3, E_4$  be smooth vector bundles over  $M$  having covariant derivatives  $\nabla^{E_1}, \nabla^{E_2}, \nabla^{E_3}, \nabla^{E_4}$  respectively, let  $A := E_1 \otimes_M E_2 \otimes_M E_3 \otimes_M E_4$  and  $B := E_1 \otimes_M E_3 \otimes_M E_2 \otimes_M E_4$ , and let  $\nabla^A$  and  $\nabla^B$  denote the induced covariant derivatives.*

*If  $(23) \in \Gamma(A^* \otimes_M B)$  denotes the tensor field which maps  $e_1 \otimes_M e_2 \otimes_M e_3 \otimes_M e_4 \in A$  to  $e_1 \otimes_M e_3 \otimes_M e_2 \otimes_M e_4 \in B$  (i.e.  $(23)$  transposes the second and third factors), then  $(23)$  is a parallel tensor field with respect to the covariant derivative induced on the vector bundle  $A^* \otimes_M B \rightarrow M$ , i.e.  $\nabla^{A^* \otimes_M B} (23) = 0$ .*

*Proof.* Let  $X \in \Gamma(TM)$ . Then

$$\begin{aligned} & (e_1 \otimes_M e_2 \otimes_M e_3 \otimes_M e_4) \cdot_{A^*} \nabla_X^{A^* \otimes_M B} (23) \\ &= \nabla_X^B ((e_1 \otimes_M e_2 \otimes_M e_3 \otimes_M e_4) \cdot_{A^*} (23)) - \nabla_X^A (e_1 \otimes_M e_2 \otimes_M e_3 \otimes_M e_4) \cdot_{A^*} (23) \\ &= \nabla_X^B (e_1 \otimes_M e_3 \otimes_M e_2 \otimes_M e_4) \\ &\quad - \nabla_X^{E_1} e_1 \otimes_M e_3 \otimes_M e_2 \otimes_M e_4 \\ &\quad - e_1 \otimes_M \nabla_X^{E_3} e_3 \otimes_M e_2 \otimes_M e_4 \\ &\quad - e_1 \otimes_M e_3 \otimes_M \nabla_X^{E_2} e_2 \otimes_M e_4 \\ &\quad - e_1 \otimes_M e_3 \otimes_M e_2 \otimes_M \nabla_X^{E_4} e_4 \\ &= \nabla_X^B (e_1 \otimes_M e_3 \otimes_M e_2 \otimes_M e_4) - \nabla_X^B (e_1 \otimes_M e_3 \otimes_M e_2 \otimes_M e_4) \\ &= 0. \end{aligned}$$

Because  $X$  is arbitrary, this shows that  $(e_1 \otimes_M e_2 \otimes_M e_3 \otimes_M e_4) \cdot_{A^*} \nabla^{A^* \otimes_M B} (23) = 0$ . This extends linearly to general tensors, so  $\nabla^{A^* \otimes_M B} (23) = 0$ , as desired.  $\square$

The fact that all transposition tensor fields are parallel implies that all permutation tensor fields are parallel, since every permutation is just the product of transpositions. This gives as an easy corollary that a covariant derivative operation commutes with a permutation operation, which has quite a succinct statement using the permutation superscript notation.

**Corollary 8.12** (Permutation tensor fields are parallel). *Let  $E_1, \dots, E_k$  be smooth vector bundles over  $M$  each having a covariant derivative, and let  $A := E_1 \otimes_M \dots \otimes_M E_k$  and  $B := E_{\sigma^{-1}(1)} \otimes_M \dots \otimes_M E_{\sigma^{-1}(k)}$ . If  $\sigma \in S_k$  is interpreted as the tensor field in  $\Gamma(A^* \otimes_M B)$  which maps  $e_1 \otimes_M \dots \otimes_M e_k$  to  $e_{\sigma^{-1}(1)} \otimes_M \dots \otimes_M e_{\sigma^{-1}(k)}$ , then  $\sigma$  is a parallel tensor field. Stated using the superscript notation, with  $X \in \Gamma(TM)$  and  $a \in \Gamma(A)$ ,*

$$\nabla_X^B a^\sigma = (\nabla_X^A a)^\sigma.$$

*Proof.* This follows from the fact that  $\sigma$  can be written as the product of transpositions;  $\nabla_X \sigma = 0$  because of the product rule and because each transposition is parallel. The claim regarding commutation with the superscript permutation follows easily from its definition.

$$\nabla_X^B a^\sigma = \nabla_X^B (a \cdot_{A^*} \sigma) = a \cdot_{A^*} \nabla_X^{A^* \otimes_M B} \sigma + \nabla_X^A a \cdot_{A^*} \sigma = (\nabla_X^A a)^\sigma,$$

using the fact that  $\nabla_X^{A^* \otimes_M B} \sigma = 0$ , since  $\sigma$  is a parallel tensor field.  $\square$

## 9 Decomposition of $\pi_E^{TE}: TE \rightarrow E$

In using the calculus of variations on a manifold  $M$  where the Lagrangian is a function of  $TM$  (this form of Lagrangian is ubiquitous in mechanics), taking the first variation involves passing to  $TTM$ . Without a way to decompose variations into more tractable components, the standard integration-by-parts trick [6, pg. 16] can't be applied. The notion of a local trivialization of  $TTM$  via choice of coordinates on  $M$  is one way to provide such a decomposition. A coordinate chart  $(U, \phi: U \rightarrow \mathbb{R}^n)$  on  $M$  establishes a locally trivializing diffeomorphism  $TTU \cong \phi(U) \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ . However such a trivialization imposes an artificial additive structure on  $TTU$  depending on the [non-canonical] choice of coordinates, only gives a local formulation of the relevant objects, and the ensuing coordinate calculations don't give clear insight into the geometric structure of the problem. The notion of the linear connection remedies this.

A **linear connection** on the vector bundle  $\pi: E \rightarrow M$  is a subbundle  $H \rightarrow E$  of  $\pi_E^{TE}: TE \rightarrow E$  such that  $TE = H \oplus_E VE$  and  $T\lambda_a \cdot H_x = H_{ax}$  for all  $a \in \mathbb{R} \setminus \{0\}$  and  $x \in E$ , where  $\lambda_a: E \rightarrow E$ ,  $e \mapsto ae$  is the scalar multiplication action of  $a$  on  $E$  [8, pg. 512]. The bundle  $H \rightarrow E$  may also be called a **horizontal space** of the vector bundle  $\pi_E^{TE}: TE \rightarrow E$  (“a” is used instead of “the” because a choice of  $H \rightarrow E$  is generally non-unique). For convenience, define  $h := \nabla \pi \in \Gamma(\pi^* TM \otimes_E T^* E)$ , noting then that  $VE = \ker h$ .

A linear connection can equivalently be specified by what is known as a connection map; essentially a projection onto the vertical bundle. This is a slightly more active formulation than just the specification of a horizontal space, as a covariant derivative can be defined directly in terms of the connection map – see [8, pg. 518], [4, pg. 128], [5, pg. 173], and [12, pg. 208].

**Proposition 9.1** (Connection map formulation of a linear connection). *If  $v \in \Gamma(\pi^* E \otimes_E T^* E)$  (i.e.  $v: TE \rightarrow E$  is a smooth vector bundle morphism over  $\pi$ ) is a left-inverse for  $\iota_{VE}^{*E} \in \Gamma(VE \otimes_E \pi^* E^*)$  that is equivariant with respect to  $T\lambda_a$  and  $\lambda_a$  (i.e.  $v \cdot T\lambda_a = \pi^* \lambda_a \cdot v$ ) [12, pg. 245], then  $H := \ker v \leq TE$  defines a linear connection on the vector bundle  $\pi: E \rightarrow M$ . Such a map  $v$  is called the **connection map** associated to  $H$ . Conversely, given a linear connection  $H$ , there is exactly one connection map defining  $H$  in the stated sense.*

*Proof.* That  $v$  is a left-inverse for  $\iota_{VE}^{*E}$  implies that  $v$  has full rank, so  $H := \ker v$  defines a subbundle of  $\pi_E^{TE}: TE \rightarrow E$  having the same rank as  $TM$ . Because  $v$  is smooth,  $H$  is a smooth subbundle. Furthermore, the condition implies that  $V_e E \cap H_e = \{0\}$  for each  $e \in E$ , and therefore  $TE = H \oplus_E VE$  by a rank-counting argument.

If  $x \in TE$  and  $a \in \mathbb{R} \setminus \{0\}$ , then  $v \cdot T\lambda_a \cdot x = \pi^* \lambda_a \cdot v \cdot x$ , which equals zero if and only if  $v \cdot x = 0$ , i.e. if and only if  $x \in H$ . Thus  $T\lambda_a \cdot H = H$ . This establishes  $H \rightarrow E$  as a linear connection.

Conversely, if  $H$  is a linear connection and  $v_1$  and  $v_2$  are connection maps for  $H$ , then  $v_1 \cdot_{TE} \iota_{VE}^{*E} = \text{Id}_{\pi^* E} = v_2 \cdot_{TE} \iota_{VE}^{*E}$ . Then because the image of  $\iota_{VE}^{*E}$  is all of  $VE$ , it follows that  $v_1|_{VE} = v_2|_{VE}$ . Since  $v_1|_H = 0 = v_2|_H$  by definition, and since  $TE = H \oplus_E VE$ , this shows that  $v_1 = v_2$ . Uniqueness of

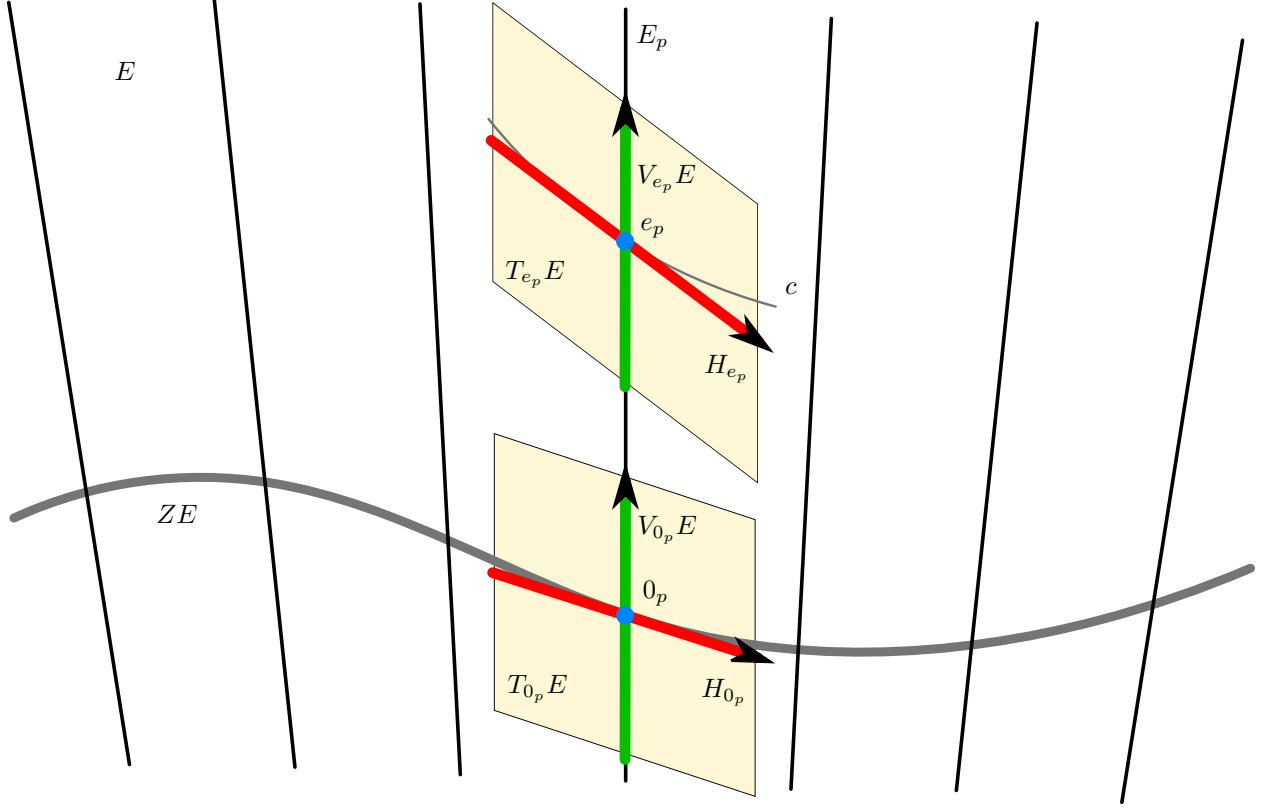


Figure 9.1: A diagram representing the decomposition of  $TE \rightarrow E$  into horizontal and vertical subbundles. The vertical lines represent individual fibers of  $E$ , while  $p \in M$ ,  $e_p \in E_p$ ,  $0_p \in E_p$  denotes the zero vector of  $E_p$ , and  $ZE$  denotes the zero subbundle of  $E$ ;  $ZE \cong M$ . By the equivariance property of the linear connection,  $ZE$  is a submanifold of  $E$  which is entirely horizontal (its tangent space is entirely composed of horizontal vectors). The tangent spaces  $T_{0_p}E$  and  $T_{e_p}E$  are drawn; green arrows representing the vertical subspaces (“along” the fibers), red arrows representing the horizontal subspaces. Finally,  $c$  is a horizontal curve passing through  $e_p$ .

connection maps has been established. To show existence, define  $v := \iota_{\pi^*E}^{VE} \cdot v_E \text{pr}_{VE} \in \Gamma(\pi^*E \otimes_E T^*E)$ , where  $\text{pr}_{VE}: H \oplus_E VE \rightarrow VE$  be the canonical projection, recalling that  $H \oplus_E VE = TE$ . It is easily shown that  $v$  is a connection map for  $H$ .  $\square$

**Proposition 9.2** (Decomposing  $\pi_E^{TE}: TE \rightarrow E$ ). *If  $v \in \Gamma(\pi^*E \otimes_E T^*E)$  is a connection map, then*

$$h \oplus_E v: TE \rightarrow \pi^*TM \oplus_E \pi^*E \quad (9.1)$$

*is a smooth vector bundle isomorphism over  $\text{Id}_E$ . See Figure 9.1.*

*Proof.* Because  $TE = H \oplus_E VE$ , and  $H = \ker v$  and  $VE = \ker h$ , the fiber-wise restriction

$$h \oplus_E v|_{T_eE}: T_eE \rightarrow (\pi^*TM \oplus_E \pi^*E)_e \cong T_{\pi(e)}M \oplus E_{\pi(e)}$$

is a linear isomorphism for each  $e \in E$ . The map is a smooth vector bundle morphism over  $\text{Id}_E$  by construction. It is therefore a smooth vector bundle isomorphism over  $\text{Id}_E$ .  $\square$

*Remark 9.3* (Linear connection/covariant derivative correspondence). Given a covariant derivative  $\nabla^E$  on a smooth vector bundle  $\pi: E \rightarrow M$ , there is a naturally induced linear connection, defined via the connection map

$$\begin{aligned} v: TE &\rightarrow E, \\ \delta_\epsilon \Theta &\mapsto \nabla_{\delta_\epsilon}^{(\pi \circ \Theta)^* E} \Theta, \end{aligned} \tag{9.2}$$

where  $\Theta: I \rightarrow E$  is a variation of  $\theta \in E$ . Here,  $\nabla^{(\pi \circ \Theta)^* E}$  denotes the pullback of the covariant derivative  $\nabla^E$  through the map  $\pi \circ \Theta$  (see (8.7)). Conceptually, all  $v$  does is replace an ordinary derivative ( $\delta_\epsilon$ ) with the corresponding covariant one ( $\nabla_{\delta_\epsilon}^{(\pi \circ \Theta)^* E}$ ).

Conversely, given a connection map  $v \in \Gamma(\pi^* E \otimes_E T^* E)$  for a linear connection  $H \rightarrow E$ , there is a naturally induced covariant derivative  $\nabla^E$  on the smooth vector bundle  $\pi: E \rightarrow M$ , defined by

$$\begin{aligned} \nabla^E: \Gamma(E) &\rightarrow \Gamma(E \otimes_M T^* M), \\ \sigma &\mapsto \sigma^* v \cdot_{\sigma^* TE} \nabla^{M \rightarrow E} \sigma. \end{aligned}$$

The scaling equivariance of  $v$  is critical for showing that this map actually defines a covariant derivative. Full type safety should be observed here; by the contravariance of the pullback of bundles (see (6.3)),  $\sigma^* \pi^* E \cong (\pi \circ \sigma)^* E = \text{Id}_M^* E \cong E$ , so

$$\sigma^* v \in \Gamma(\sigma^* (\pi^* E \otimes_E T^* E)) \cong \Gamma(\sigma^* \pi^* E \otimes_M \sigma^* T^* E) \cong \Gamma(E \otimes_M \sigma^* T^* E),$$

and therefore  $\sigma^* v \cdot \nabla \sigma \in \Gamma(E \otimes_M T^* M)$  as desired. This connection map construction of a covariant derivative gives (8.7) as an immediate consequence via the chain rule for the tangent map.

The following construction is an abstraction of taking partial derivatives of a function, inspired by [11, pg. 277]. Instead of taking partial derivatives with respect to individual coordinates, partial covariant derivatives along distributions over the base manifold are formed, where the distributions (subbundles) decompose the base manifold's tangent bundle into a direct sum. Such a construction conveniently captures the geometry of maps with respect to the geometry of its domain.

**Proposition 9.4** (Partial covariant derivatives). *Let  $L \in C^\infty(M, \mathbb{R})$ , and for each  $i \in \{1, \dots, n\}$  let  $F_i \rightarrow M$  be a smooth vector bundle. If, for each  $i \in \{1, \dots, n\}$ ,  $c_i \in \Gamma(F_i \otimes_M T^* M)$  such that  $c_1 \oplus_M \dots \oplus_M c_n \in \Gamma((F_1 \oplus_M \dots \oplus_M F_n) \otimes_M T^* M)$  is a smooth vector bundle isomorphism over  $\text{Id}_E$ , then there exist unique sections  $L_{c_i} \in \Gamma(F_i^*)$  for each  $i \in \{1, \dots, n\}$  such that*

$$\nabla^{M \rightarrow \mathbb{R}} L = L_{c_1} \cdot_{F_1} c_1 + \dots + L_{c_n} \cdot_{F_n} c_n.$$

*This decomposition of  $\nabla L$  provides what will be called **partial covariant derivatives** of  $L$  (with respect to the given decomposition).*

*Proof.* The following equivalences provide a formula for directly defining  $L_{c_1}, \dots, L_{c_n}$ .

$$\begin{aligned} \nabla L &= L_{c_1} \cdot_{F_1} c_1 + \dots + L_{c_n} \cdot_{F_n} c_n \\ \iff \nabla L &= (L_{c_1} \oplus_M \dots \oplus_M L_{c_n}) \cdot_{F_1 \oplus_M \dots \oplus_M F_n} (c_1 \oplus_M \dots \oplus_M c_n) \\ \iff \nabla L \cdot_{TM} (c_1 \oplus_M \dots \oplus_M c_n)^{-1} &= L_{c_1} \oplus_M \dots \oplus_M L_{c_n}. \end{aligned}$$

Existence and uniqueness is therefore proven.  $\square$

**Corollary 9.5** (Horizontal/vertical derivatives). *Let  $h := \nabla \pi \in \Gamma(\pi^* TM \otimes_E T^* E)$  as before. If  $v \in \Gamma(\pi^* E \otimes_E T^* E)$  is a connection map, and if  $L: E \rightarrow \mathbb{R}$  is smooth, then there exist unique  $L_h \in \Gamma(\pi^* T^* M)$  and  $L_v \in \Gamma(\pi^* E^*)$  such that  $\nabla L = L_h \cdot_{\pi^* TM} h + L_v \cdot_{\pi^* E} v$ .*

It should be noted that the basepoint-preserving issue discussed in Section 7 plays a role in choosing to use the tensor field formulation of  $h: TE \rightarrow TM$  and  $v: TE \rightarrow E$ . In particular, without preserving the basepoint (via the  $\pi$ -pullback of  $TM$  and  $E$  to form  $h \in \Gamma(\pi^*TM \otimes_E T^*E)$  and  $v \in \Gamma(\pi^*E \otimes_E T^*E)$ ), the map  $h \oplus_E v$  would not be a smooth bundle isomorphism, and the horizontal and vertical derivatives would be maps of the form  $L_{,h}: E \rightarrow T^*M$  and  $L_{,v}: E \rightarrow E^*$ , but that, critically, are *not* sections of smooth vector bundles, and can only claim to be smooth [fiber] bundle morphisms. Derivative trivializations will be central in calculating the first and second variations of an energy functional having Lagrangian  $L$  (see (12.1) and (13.1)).

## 10 Curvature and Commutation of Derivatives

A ubiquitous consideration in mathematics is to determine when two operations commute. In the setting of tensor calculus, this often manifests itself in determining the commutativity (or lack thereof) of two covariant derivatives. Here, “covariant derivatives” may refer to both linear covariant derivatives and the tangent map operator (see (7.1)). This unified categorization of derivatives will now be leveraged to show that certain fiber bundles are flat (in a sense analogous to the vanishing of a curvature endomorphism) with respect to particular covariant derivatives. This reduces the work often done showing commutativity of derivatives in the derivation of the first variation of a function in the calculus of variations to the simple statement that a particular tensor field is symmetric, which is comes as a corollary to the aforementioned flatness.

In this section, the symbol  $\nabla$  may denote  $\nabla$  or  $\nabla$ , depending on context. This eases the expression of repeated covariant derivatives, such as the covariant Hessian of a section (see below), and is an example of telescoping notation as discussed in Section 3.

If  $\pi: E \rightarrow M$  defines a smooth [fiber] bundle whose space of sections  $\Gamma(E)$  has two repeated covariant derivatives defined and if  $\nabla^{TM}$  is a symmetric linear covariant derivative (meaning  $\nabla_X Y - \nabla_Y X = [X, Y]$  for  $X, Y \in \Gamma(TM)$ ), then the tensor contraction

$$\nabla^2 \sigma : (X \otimes_M Y - Y \otimes_M X)$$

is an expression measuring the non-commutativity of the  $X$  and  $Y$  derivatives of  $\sigma$ . The quantity  $\nabla^2 \sigma$  will be called the **covariant Hessian** of  $\sigma$ , because it generalizes the Hessian of elementary calculus; it contains only second-derivative information, and in the special case seen below, it is symmetric in the argument components. It should be noted that if  $F \rightarrow M$  is the vector bundle such that  $\nabla \sigma \in \Gamma(F \otimes_M T^*M)$ , then  $\nabla^2 \sigma \in \Gamma(F \otimes_M T^*M \otimes_M T^*M)$ . Intentionally leaving the  $\nabla$  and  $\cdot$  symbols undecorated in preference of contextual interpretation, unwinding the expression above gives

$$\begin{aligned} \nabla^2 \sigma : (X \otimes_M Y - Y \otimes_M X) &= \nabla_Y \nabla_X \sigma \cdot X - \nabla_X \nabla_Y \sigma \cdot Y \\ &= \nabla_Y \nabla_X \sigma - \nabla \sigma \cdot \nabla_Y X - \nabla_X \nabla_Y \sigma + \nabla \sigma \cdot \nabla_X Y \\ &= -\nabla_X \nabla_Y \sigma + \nabla_Y \nabla_X \sigma + \nabla \sigma \cdot [X, Y] \\ &= -\nabla_X \nabla_Y \sigma + \nabla_Y \nabla_X \sigma + \nabla_{[X, Y]} \sigma, \end{aligned}$$

which is syntactically identical to the common definition for the [Riemannian] curvature endomorphism  $R(X, Y) \sigma$ . In the traditional setting, where  $\nabla^E$  is a linear covariant derivative on vector bundle  $E$ , the curvature endomorphism takes the form of a tensor field  $R^E \in \Gamma(E \otimes_M E^* \otimes_M T^*M \otimes_M T^*M)$ . In this setting however, because  $\nabla^E$  may be nonlinear, such a tensorial formulation doesn’t generally exist. Instead,

$$R^E(X, Y) := -\nabla_X \nabla_Y^E + \nabla_Y \nabla_X^E + \nabla_{[X, Y]}^E$$

defines a second-order covariant differential operator (“covariant” meaning tensorial in the  $X$  and  $Y$  components). Put differently,

$$R^E(X, Y) \sigma = \nabla^2 \sigma : (X \otimes_M Y - Y \otimes_M X),$$



which will be called the (possibly nonlinear) **curvature operator**, measures the non-commutativity of the  $X$  and  $Y$  derivatives of  $\sigma$ . If  $R^E$  is identically zero, then the bundle  $E$  is said to be **flat** with respect to the relevant connections/covariant derivatives.

There are two particularly important instances of flat bundles. The first is the trivial line bundle defined by  $\pi^{\mathbb{R} \times S}$  (whose space of smooth sections, as discussed in Section 4, is naturally identified with  $C^\infty(S, \mathbb{R})$ ). In this case,  $\nabla^{S \rightarrow \mathbb{R}} f \in \Gamma(T^*S)$ , and  $\nabla^2 f \equiv \nabla^{T^*S} \nabla^{S \rightarrow \mathbb{R}} f \in \Gamma(T^*S \otimes_S T^*S)$  is the object referred to in most literature as the covariant Hessian of  $f$ . Here,  $R^{S \rightarrow \mathbb{R}}(X, Y)f$  is a real-valued function on  $S$ .

**Proposition 10.1** (Symmetry of covariant Hessian on functions). *Let  $S$  be a smooth manifold and let  $\nabla^{TS}$  be a symmetric covariant derivative. If  $f \in C^\infty(S, \mathbb{R})$ , then  $\nabla^2 f \in \Gamma(T^*S \otimes_S T^*S)$  is a symmetric tensor field (i.e. it has a (12) symmetry). Here, the covariant derivative on  $C^\infty(S, \mathbb{R})$  is  $\nabla^{S \rightarrow \mathbb{R}}$  as defined above.*

*Proof.* Let  $X, Y \in \Gamma(TS)$ . Recall that  $\nabla f \equiv df \in \Gamma(T^*S)$ . Then

$$\begin{aligned} \nabla^2 f &: (X \otimes_S Y - Y \otimes_S X) \\ &= \nabla_Y \nabla f \cdot X - \nabla_X \nabla f \cdot Y \\ &= \nabla_Y (\nabla f \cdot X) - \nabla f \cdot \nabla_Y X - \nabla_X (\nabla f \cdot Y) + \nabla f \cdot \nabla_X Y \\ &= \nabla (\nabla f \cdot X) \cdot Y - \nabla (\nabla f \cdot Y) \cdot X + \nabla f \cdot [X, Y] && \text{(by symmetry of } \nabla^{TS} \text{)} \\ &= -\nabla f \cdot [X, Y] + \nabla f \cdot [X, Y] && \text{(by definition of } [X, Y] \text{)} \\ &= 0. \end{aligned}$$

Because  $X \otimes_S Y$  is pointwise-arbitrary in  $TS \otimes_S TS$ , this shows that  $\nabla^2 f$  is symmetric. Equivalently stated,  $R^{S \rightarrow \mathbb{R}}$  is identically zero, and therefore the relevant bundle is flat.  $\square$

The second important case involves the nonlinear covariant derivative  $\nabla^{M \rightarrow S}$  on  $C^\infty(M, S)$ . Here, if  $\phi \in C^\infty(M, S)$ , then

$$\nabla^2 \phi \equiv \nabla^{\phi^* TS \otimes_M T^* M} \nabla^{M \rightarrow S} \phi \in \Gamma(\phi^* TS \otimes_M T^* M \otimes_M T^* M),$$

so  $R^{M \rightarrow S}(X, Y)\phi \in \Gamma(\phi^* TS)$ .

**Proposition 10.2** (Symmetry of covariant Hessian on maps). *Let  $M$  and  $S$  be smooth manifolds and let  $\nabla^{TM}$  and  $\nabla^{TS}$  be symmetric covariant derivatives. If  $\phi \in C^\infty(M, S)$ , then  $\nabla^2 \phi \in \Gamma(\phi^* TS \otimes_M T^* M \otimes_M T^* M)$  is a tensor field which is symmetric in the two  $T^* M$  components (i.e. it has a (23) symmetry). Here, the covariant derivative on  $C^\infty(M, S)$  is  $\nabla^{M \rightarrow S}$  as defined above.*

*Proof.* Let  $X, Y \in \Gamma(TM)$  and  $f \in C^\infty(S, \mathbb{R})$ , so that  $\phi^* \nabla f \in \Gamma(\phi^* TS)$ . Then

$$\begin{aligned} \phi^* \nabla f \cdot \phi^*_{TS} R^{M \rightarrow S}(X, Y)\phi &= \phi^* \nabla f \cdot (-\nabla_X \nabla_Y \phi + \nabla_Y \nabla_X \phi + \nabla_{[X, Y]}\phi) \\ &= -\nabla_X (\phi^* \nabla f \cdot \nabla \phi \cdot Y) + \nabla_X \phi^* \nabla f \cdot \nabla \phi \cdot Y \\ &\quad + \nabla_Y (\phi^* \nabla f \cdot \nabla \phi \cdot X) - \nabla_Y \phi^* \nabla f \cdot \nabla \phi \cdot X \\ &\quad + \phi^* \nabla f \cdot \nabla \phi \cdot [X, Y] \\ &= -\nabla_X (\nabla \phi^* f \cdot Y) + \nabla_Y (\nabla \phi^* f \cdot X) + \nabla \phi^* f \cdot [X, Y] \\ &\quad + (\phi^* \nabla^2 f \cdot \nabla \phi \cdot X) \cdot \nabla \phi \cdot Y - (\phi^* \nabla^2 f \cdot \nabla \phi \cdot Y) \cdot \nabla \phi \cdot X \\ &= -\nabla (\nabla \phi^* f \cdot Y) \cdot X + \nabla (\nabla \phi^* f \cdot X) \cdot Y + \nabla \phi^* f \cdot [X, Y] \\ &\quad - \phi^* \nabla^2 f : ((\nabla \phi \cdot X) \otimes_M (\nabla \phi \cdot Y) - (\nabla \phi \cdot Y) \otimes_M (\nabla \phi \cdot X)). \end{aligned}$$

By definition,  $-\nabla (\nabla \phi^* f \cdot Y) \cdot X + \nabla (\nabla \phi^* f \cdot X) \cdot Y = -\nabla \phi^* f \cdot [X, Y]$ , which cancels out the other term. By (10.1),  $\nabla^2 f$  is symmetric, so the final term is zero. Because  $\phi^* \nabla f$  is pointwise-arbitrary in  $\phi^* T^* S$  and  $X$  and  $Y$  are pointwise-arbitrary in  $TM$ , this shows that  $R^{M \rightarrow S}$  is identically zero, so the bundle defined by  $\pi_M^{S \times M}: S \times M \rightarrow M$ , whose space of sections is identified with  $C^\infty(M, S)$ , is flat, and therefore  $\nabla^2 \phi$  is symmetric in its two  $T^* M$  components.  $\square$

The construction used in (9.4) can be applied to nonlinear as well as linear covariant derivatives to considerable advantage. For example, if  $\psi: M \times N \rightarrow L$ , where  $M, N, L$  are smooth manifolds and  $p_M := \text{pr}_1^{M \times N}$  and  $p_N := \text{pr}_2^{M \times N}$ , then define  $\psi_{,M} \in \Gamma(\psi^*TL \otimes_{M \times N} p_M^*T^*M)$  and  $\psi_{,N} \in \Gamma(\psi^*TL \otimes_{M \times N} p_N^*T^*N)$  by

$$\nabla \psi = \nabla^{M \times N \rightarrow L} \psi = \psi_{,M} \cdot p_M^*TM \nabla p_M + \psi_{,N} \cdot p_N^*TN \nabla p_N.$$

This gives a convenient way to express partial covariant derivatives, which will be used heavily in Part II in calculating the first and second variations of an energy functional. Note that in this parlance,  $\psi_{,(M \times N)}$  is the full tangent map  $\nabla \psi$ .

Defining second partial covariant derivatives  $\psi_{,MM}$ ,  $\psi_{,MN}$ ,  $\psi_{,NM}$  and  $\psi_{,NN}$  by

$$\begin{aligned} \nabla \psi_{,M} &= \psi_{,MM} \cdot \nabla p_M + \psi_{,MN} \cdot \nabla p_N \text{ and} \\ \nabla \psi_{,N} &= \psi_{,NM} \cdot \nabla p_M + \psi_{,NN} \cdot \nabla p_N, \end{aligned}$$

the symmetry of the covariant Hessian of  $\psi$  can be used to show various symmetries these second derivatives.

**Proposition 10.3** (Symmetries of partial covariant derivatives). *With  $\psi$  and its second partial covariant derivatives as above,*

$$\psi_{,MM} \in \Gamma(\psi^*TL \otimes_{M \times N} p_M^*T^*M \otimes_{M \times N} p_M^*T^*M)$$

*and  $\psi_{,NN}$  (having analogous type) are (23)-symmetric (i.e.  $(\psi_{,MM})^{(23)} = \psi_{,MM}$  and  $(\psi_{,NN})^{(23)} = \psi_{,NN}$ ) and the mixed, second partial covariant derivatives*

$$\begin{aligned} \psi_{,MN} &\in \Gamma(\psi^*TL \otimes_{M \times N} p_M^*T^*N \otimes_{M \times N} p_N^*T^*N) \text{ and} \\ \psi_{,NM} &\in \Gamma(\psi^*TL \otimes_{M \times N} p_N^*T^*N \otimes_{M \times N} p_M^*T^*M) \end{aligned}$$

*are mutually (23)-symmetric (i.e.  $\psi_{,MN} = (\psi_{,NM})^{(23)}$ ).*

*Proof.* Let  $X, Y \in \Gamma(TM \oplus TN)$ . If  $TP_N \cdot X = 0$  and  $TP_M \cdot Y = 0$ , then

$$\begin{aligned} 0 &= \nabla^2 \psi : (X \otimes_{M \times N} Y - Y \otimes_{M \times N} X) \text{ (by (10.2))} \\ &= \psi_{,MM} : (\nabla p_M \cdot X \otimes_{M \times N} \nabla p_M \cdot Y - \nabla p_M \cdot Y \otimes_{M \times N} \nabla p_M \cdot X) \\ &\quad + \psi_{,MN} : (\nabla p_M \cdot X \otimes_{M \times N} \nabla p_N \cdot Y - \nabla p_M \cdot Y \otimes_{M \times N} \nabla p_N \cdot X) \\ &\quad + \psi_{,NM} : (\nabla p_N \cdot X \otimes_{M \times N} \nabla p_M \cdot Y - \nabla p_N \cdot Y \otimes_{M \times N} \nabla p_M \cdot X) \\ &\quad + \psi_{,NN} : (\nabla p_N \cdot X \otimes_{M \times N} \nabla p_N \cdot Y - \nabla p_N \cdot Y \otimes_{M \times N} \nabla p_N \cdot X) \\ &= \psi_{,MN} : (\nabla p_M \cdot X \otimes_{M \times N} \nabla p_N \cdot Y - \nabla p_N \cdot Y \otimes_{M \times N} \nabla p_M \cdot X) \\ &= (\psi_{,MN} - (\psi_{,NM})^{(23)}) : (\nabla p_M \cdot X \otimes_{M \times N} \nabla p_N \cdot Y). \end{aligned}$$

Because  $\nabla p_M \cdot X$  and  $\nabla p_N \cdot Y$  are pointwise-arbitrary in  $p_M^*TM$  and  $p_N^*TN$  respectively, this implies that  $\psi_{,MN} = (\psi_{,NM})^{(23)}$ . Analogous calculations (setting  $\nabla p_M \cdot X = 0$  and  $\nabla p_M \cdot Y = 0$  and then separately setting  $\nabla p_N \cdot X = 0$  and  $\nabla p_N \cdot Y = 0$ ) show that  $\psi_{,MM} = (\psi_{,MM})^{(23)}$  and  $\psi_{,NN} = (\psi_{,NN})^{(23)}$ .  $\square$

There are two final results regarding the second covariant derivative that will be especially useful in the calculation of the first and second variations of an energy functional (see (12.1) and (13.2)).

**Proposition 10.4** (Chain rule for covariant Hessian). *Let  $\pi: E \rightarrow N$  define a bundle having a first and second covariant derivative (i.e. a section of  $E$  can be covariantly differentiated twice). If  $\phi: M \rightarrow N$  and  $e \in \Gamma(E)$ , then*

$$\nabla^2 \phi^* e = \phi^* \nabla^2 e :_{\phi^*TN} (\nabla \phi \boxtimes_M \nabla \phi) + \phi^* \nabla e \cdot_{\phi^*TN} \nabla \nabla \phi.$$

*Proof.* Let  $X \in \Gamma(TM)$ . Then

$$\begin{aligned}
\nabla^2 \phi^* e \cdot X &= \nabla_X \nabla \phi^* E \phi^* e \\
&= \nabla_X (\phi^* \nabla^E e \cdot \phi^* TN \nabla \phi) \\
&= \nabla_X (\phi^* \nabla e) \cdot \phi^* TN \nabla \phi + \phi^* \nabla e \cdot \phi^* TN \nabla_X \nabla \phi \\
&= (\phi^* \nabla^2 e \cdot \phi^* TN \nabla \phi \cdot X) \cdot \phi^* TN \nabla \phi + \phi^* \nabla e \cdot \phi^* TN \nabla_X \nabla \phi \\
&= [\phi^* \nabla^2 e \cdot \phi^* TN (\nabla \phi \boxtimes_M \nabla \phi) + \phi^* \nabla e \cdot \phi^* TN \nabla \nabla \phi] \cdot X.
\end{aligned}$$

Because  $X$  is pointwise-arbitrary in  $TM$ , this establishes the desired equality.  $\square$

**Proposition 10.5** (Pullback curvature endomorphism). *Let  $\pi: E \rightarrow N$  define a vector bundle having first and second covariant derivatives. If  $\phi: M \rightarrow N$ , then  $R^{\phi^* TN} = \phi^* R^{TN} \cdot \phi^* TN (\nabla \phi \boxtimes_M \nabla \phi)$ .*

*Proof.* Note that  $R^{\phi^* TN} \in \Gamma(\phi^* TN \otimes_M \phi^* T^* N \otimes_M T^* M \otimes_M T^* M)$ . Let  $X, Y \in \Gamma(TM)$  and let  $Z \in \Gamma(TN)$ , so that  $\phi^* Z \in \Gamma(\phi^* TN)$ . Then

$$\begin{aligned}
&(\text{Id}_{\phi^* TN \otimes_M \phi^* Z}) \cdot \phi^* TN \otimes_M \phi^* T^* N R^{\phi^* TN} \cdot_{TM} (X \otimes_M Y) \\
&= R^{\phi^* TN} (X, Y) (\phi^* Z) \\
&= \nabla^2 \phi^* Z \cdot_{TM} (X \wedge_M Y) \\
&= \phi^* \nabla^2 Z \cdot \phi^* TN (\nabla \phi \boxtimes_M \nabla \phi) \cdot_{TM} (X \wedge_M Y) \\
&\quad + \phi^* \nabla Z \cdot \phi^* TN \nabla \nabla \phi \cdot_{TM} (X \wedge_M Y) \tag{by (10.4)} \\
&= \phi^* \nabla^2 Z \cdot \phi^* TN ((\nabla \phi \cdot X) \wedge_M (\nabla \phi \cdot Y)) + \phi^* \nabla Z \cdot \phi^* TN 0 \tag{by symmetry of  $\nabla \nabla \phi$ } \\
&= (\phi^* ((\text{Id}_{TN} \otimes_N Z) \cdot_{TN \otimes_N T^* N} R^{TN})) \cdot \phi^* TN ((\nabla \phi \cdot X) \otimes_M (\nabla \phi \cdot Y)) \\
&= (\text{Id}_{\phi^* TN \otimes_M \phi^* Z}) \cdot \phi^* TN \otimes_M \phi^* T^* N \phi^* R^{TN} \cdot \phi^* TN (\nabla \phi \boxtimes_M \nabla \phi) \cdot_{TM} (X \otimes_M Y),
\end{aligned}$$

and because  $X, Y$  and  $\phi^* Z$  are pointwise-arbitrary in their respective spaces, this establishes the desired equality.  $\square$

A common operation is to evaluate a covariant derivative along a single tangent vector. One can express a single tangent vector as a section of a particular pullback bundle, the map being the constant map evaluating to the basepoint of the vector. This allows the richly-typed formalism of pullback bundles to be used to evaluate derivatives at a point, particularly noting that this safely deals with the overloading of the natural pairing operator  $\cdot$  (see Section 5).

**Proposition 10.6** (Evaluation commutes with non-involved derivatives). *Let  $A$  and  $B$  be smooth manifolds and let  $\sigma \in \Gamma(E)$  for some smooth bundle  $E \rightarrow A \times B$  having a covariant derivative  $\nabla^E$ . If  $b \in B$  and the map  $z: A \rightarrow A \times B$ ,  $a \mapsto (a, b)$  represents evaluation at  $b$ , then*

$$z^* (\sigma, A) = (z^* \sigma)_{,A},$$

*i.e. evaluation in  $B$  commutes with a derivative along  $A$ .*

*Proof.* Let  $X \in \Gamma(TA)$ , and let  $p_A := \text{pr}_1^{A \times B}$  and  $p_B := \text{pr}_2^{A \times B}$ . Then

$$\begin{aligned}
(z^* \sigma)_{,A} \cdot X &= \nabla^{z^* E} z^* \sigma \cdot X \\
&= z^* \nabla^E \sigma \cdot z^* (TA \oplus TB) \nabla z \cdot_{TA} X \\
&= z^* (\nabla^E \sigma \cdot_{TA \oplus TB} p_A^* X) \\
&= z^* (\sigma_{,A} \cdot_{p_A^* TA} \nabla p_A \cdot_{TA \oplus TB} p_A^* X) \tag{since  $\nabla p_B \cdot_{TA \oplus TB} p_A^* X = 0$ } \\
&= z^* \sigma_{,A} \cdot_{TA} X \tag{since  $z^* p_A^* X = (p_A \circ z)^* X = \text{Id}_A^* X = X$ },
\end{aligned}$$

and because  $X$  is pointwise-arbitrary in  $TM$ , this implies that  $z^* \sigma_{,A} = (z^* \sigma)_{,A}$  as desired.  $\square$

**Proposition 10.7.** *Let  $A, B, C$  be smooth manifolds, let  $\psi: A \times B \rightarrow C$  be smooth, let  $p_A := \text{pr}_1^{A \times B}$  and  $p_B := \text{pr}_2^{A \times B}$ , and let  $X, Y \in \Gamma(TA \oplus TB)$ . If  $\nabla p_B \cdot X = 0$  and  $\nabla p_A \cdot Y = 0$ , then*

$$\psi_{,AB} : ((\nabla p_A \cdot X) \otimes_{A \times B} (\nabla p_B \cdot Y)) = \nabla_Y^{*\psi TC} \nabla_X^{A \times B \rightarrow C} \psi.$$

*Proof.* The conditions  $\nabla p_B \cdot X = 0$  and  $\nabla p_A \cdot Y = 0$  imply that  $\nabla_Y X = 0$  in the product covariant derivative. Then since  $p_A \times_{A \times B} p_B = \text{Id}_{A \times B}$ , it follows that

$$\nabla \nabla p_A \oplus_{A \times B} \nabla \nabla p_B = \nabla (\nabla p_A \oplus_{A \times B} \nabla p_B) = \nabla \nabla (p_A \times_{A \times B} p_B) = \nabla \text{Id}_{TA \oplus TB} = 0,$$

and therefore

$$\nabla_Y (\nabla p_A \cdot X) = \nabla_Y \nabla p_A \cdot X + \nabla p_A \cdot \nabla_Y X = 0 \cdot X + \nabla p_A \cdot 0 = 0.$$

For the main calculation,

$$\begin{aligned} & \psi_{,AB} : ((\nabla p_A \cdot X) \otimes_{A \times B} (\nabla p_B \cdot Y)) \\ &= (\psi_{,AB} \cdot p_B^* TB \nabla p_B \cdot TA \oplus TB Y) \cdot p_A^* TA \nabla p_A \cdot TA \oplus TB X \\ &= \left( \nabla_Y^{\psi \times A \times B p_A} \psi_{,A} \right) \cdot p_A^* TA \nabla p_A \cdot TA \oplus TB X && (\text{since } \nabla p_A \cdot Y = 0) \\ &= \nabla_Y^\psi (\psi_{,A} \cdot \nabla p_A \cdot X) - \psi_{,A} \cdot \nabla_Y^{p_A} (\nabla p_A \cdot X) && (\text{by reverse product rule}) \\ &= \nabla_Y^{*\psi TC} \nabla_X^{A \times B \rightarrow C} \psi && (\text{since } \nabla_Y (\nabla p_A \cdot X) = 0), \end{aligned}$$

as desired. □

## Part II

# Riemannian Calculus of Variations

The use of the Calculus of Variations in the Riemannian setting to develop the geodesic equations and to study harmonic maps is quite well-established. A more general formulation is required for more specific applications, such as continuum mechanics in Riemannian manifolds. The tools developed in Part I will now be used to formulate the first and second variations and Euler-Lagrange equations of an energy functional corresponding to a first-order Lagrangian. In particular, the bundle decomposition discussed in Section 9 will be needed to employ the standard integration-by-parts trick seen in the formulation of the analogous parts of the elementary Calculus of Variations. The seemingly heavy and pedantic formalism built up thus far will now show its usefulness.

In this part, let  $(M, g)$  and  $(S, h)$  be Riemannian manifolds with  $M$  compact. Calculations will be done formally in the space  $C^\infty(M, S)$ , noting that its completion under various norms will give various Sobolev spaces of maps from  $M$  to  $S$ , which are ultimately the spaces which must be considered when finding critical points of the relevant energy functionals. See [4, 5] for details on the analytical issues. Let  $dV_g$  denote the Riemannian volume form corresponding to metric  $g$ , and let  $d\bar{V}_g$  be the induced volume form on  $\partial M$ . Let  $\iota: \partial M \rightarrow M$  be the inclusion, and let  $\nu \in \Gamma(\iota^*T^*M)$  be the unit normal covector field on  $\partial M$ . Let  $E := TS \otimes_{S \times M} T^*M$  and  $\pi := \pi_S^{TS} \otimes_{S \times M} \pi_M^{T^*M}$ , making  $\pi: E \rightarrow S \times M$  a vector bundle.

The energy functionals in this section will be assumed to have the form

$$\begin{aligned} \mathcal{L}: C^\infty(M, S) &\rightarrow \mathbb{R}, \\ \phi &\mapsto \int_M L \circ \nabla \phi \, dV_g, \end{aligned}$$

where  $L: E \rightarrow \mathbb{R}$ , referred to as the **Lagrangian** of the functional, is smooth. Here,  $\nabla \phi$  could be understood to take values either in  $E = TS \otimes_{S \times M} T^*M$  or  $\phi^*TS \otimes_M T^*M$ . In the former case, the composition  $L \circ \nabla \phi$  is literal, while in the latter case, there is an implicit conversion from  $\phi^*TS \otimes_M T^*M$  to  $TS \otimes_{S \times M} T^*M$  via a fiber projection bundle morphism (see (6.2)). Either way,  $L \circ \nabla \phi: M \rightarrow \mathbb{R}$ . Let  $\nabla^{TS}$  and  $\nabla^{TM}$  denote the respective Levi-Civita connections, which induce a covariant derivative  $\nabla^E$  on  $E$  (see (8.9)). Define the connection map  $v \in \Gamma(\pi^*E \otimes_E T^*E)$  using  $\nabla^E$  as in (9.2). For convenience, the  $S \times M$  subscript will be suppressed on the “full” tensor product defining  $E$  from here forward.

## 11 Critical Points and Variations

One of the most pertinent properties of an energy functional is its set of critical points. Often, the solution to a problem in physics will take the form of minimizing a particular energy functional. Lagrangian mechanics is the quintessential example of this. This section will deal with some of the main considerations regarding such critical points.

Because the domain of a [real-valued] functional  $\mathcal{L}$  may be a nonlinear space, the relevant first derivative is the [real-valued] differential  $d\mathcal{L}$ , which is paired with the linearized variation of a map  $\phi \in C^\infty(M, S)$ . In particular, a **one-parameter variation** of  $\phi$  is a smooth map  $\Phi: M \times I \rightarrow S$ , where the  $I$  component is the variational parameter. Letting  $i$  denote the standard coordinate on  $I$ , the linearized variation is then  $\delta_i \Phi: M \rightarrow TS$ , recalling that  $\delta_i := \frac{\partial}{\partial i} |_{i=0}$ . Because  $\pi_S^{TS} \circ \delta_i \Phi = \phi$ , it follows that  $\delta_i \Phi \in \Gamma(\phi^*TS)$ , i.e.  $\delta_i \Phi$  is a vector field along  $\phi$ . The object  $\delta_i \Phi$  will be called a **linearized variation**. Call the elements of  $\Gamma(\phi^*TS)$  **linear variations**.

**Proposition 11.1** (Each linear variation is a linearized variation). *Let  $\exp: U \rightarrow S$  denote the exponential map associated to  $\nabla^{TS}$ , where  $U \subseteq TS$  is a neighborhood of the zero bundle in  $TS$  on which  $\exp$  is defined, and let  $\lambda: TS \times \mathbb{R} \rightarrow TS$ ,  $(s, \epsilon) \mapsto \epsilon s$  denote the scalar multiplication structure on  $TS$ . If  $A \in \Gamma(\phi^*TS)$  and*

if  $\Phi: U \rightarrow S$  is defined by  $\Phi := \exp \circ \lambda \circ (A \times \text{Id}_I) |_U$ , then  $\delta_i \Phi = A$ . In other words, every vector field over  $\phi$  is realized as the linearization of a one-parameter variation of  $\phi$ .

*Proof.* The map  $\Phi$  is well-defined and smooth by construction. Let  $p \in M$ . Then

$$\begin{aligned} (\delta_i \Phi)(p) &= \delta_i (\Phi(p, i)) \\ &= \delta_i (\exp \circ \lambda \circ (A \times \text{Id}_I)(p, i)) \\ &= \nabla \exp \cdot \delta_i (\lambda(A(p), i)) \\ &= \nabla \exp \cdot \delta_i (iA(p)) \\ &= \nabla \exp \cdot \left( \iota_{V_E}^{\pi^* E} |_{Z(\pi^* E)} \right) \cdot A(p) \\ &= A(p), \end{aligned}$$

where  $Z(\pi^* E)$  denotes the zero subbundle of  $\pi^* E$ . The last equality follows from a naturality property of the exponential map [10, pg. 523].  $\square$

Thus each linear variation is a linearized variation, establishing a natural identification of  $T_\phi(C^\infty(M, S))$  with  $\Gamma(\phi^* TS)$ , which will be useful when calculating the differential of a functional on  $C^\infty(M, S)$ . In fact, the exponential map construction in (11.1) is a way to construct charts for the infinite dimensional manifold  $C^\infty(M, S)$  [5, Theorem 5.2].

## 12 First Variation

This section is devoted to calculating the first variation of the previously defined energy functional. Here is where the full richness of the type system of the objects developed earlier in the paper will really show their power (and arguably, necessity). While the type-specifying notation may appear overly decorated and pedantic, subtle usage errors can be detected and avoided by keeping track of the myriad of types of the relevant objects through the sub/superscripts on covariant derivatives and natural pairings; extremely complex constructions can be made and navigated without much trouble. By contrast, performing the ensuing calculations in coordinate trivializations would result in an intractable proliferation of Christoffel symbols and indexed expressions which would prove difficult to read and would be highly prone to error.

Because the Lagrangian  $L: E \rightarrow \mathbb{R}$  is defined on a vector bundle  $\pi: E \rightarrow S \times M$  over the product space  $S \times M$ , the decomposition in (9.5) can be slightly refined. The projection  $\pi$  can be decomposed into the factors  $\pi_S := \text{pr}_S^{S \times M} \circ \pi$  and  $\pi_M := \text{pr}_M^{S \times M} \circ \pi$ , so that  $\pi = \pi_S \times_E \pi_M$ . Then  $h = \nabla \pi = \nabla \pi_S \oplus_E \nabla \pi_M$ . Let

$$\sigma := \nabla \pi_S \in \Gamma(\pi_S^* TS \otimes_E T^* E) \text{ and } \mu := \nabla \pi_M \in \Gamma(\pi_M^* TM \otimes_E T^* E).$$

The letters *sigma* and *mu* have been chosen to reflect the fact that  $L_{,\sigma} \in \Gamma(\pi_S^* T^* S)$  and  $L_{,\mu} \in \Gamma(\pi_M^* T^* M)$  give the “*S* component” (spatial) and “*M* component” (material) of the derivative  $\nabla^{E \rightarrow \mathbb{R}} L \in \Gamma(T^* E)$ . The connection map  $v$  will be retained as is, giving  $L_{,v} \in \Gamma(\pi^* E^*)$ , the “*E* component” (fiber) of  $\nabla^{E \rightarrow \mathbb{R}} L$ . See (12.5) for a discussion of how the quantities  $L_{,\mu}, L_{,\sigma}, L_{,v}$  generalize the analogous structures in the elementary treatment of the calculus of variations.

Because a one-parameter variation of  $\phi \in C^\infty(M, S)$  has the form  $\Phi: M \times I \rightarrow S$  but the energy functional  $\mathcal{L}$  involves only the *M* derivative of its argument, the partial tangent map must be used here. For the purposes of calculating the first and second variations,  $\mathcal{L}$  must be written as

$$\mathcal{L}(\phi) := \int_M L \circ \phi_{,M} dV_g.$$

**Theorem 12.1** (First variation of  $\mathcal{L}$ ). *Let  $\mathcal{L}, L, \sigma, \mu, v$  and  $\nu$  all be defined as above. If  $\phi \in C^\infty(M, S)$  and  $A \in \Gamma(\phi^* TS)$ , then*

$$d\mathcal{L}(\phi) \cdot A = \int_M A \cdot_{\phi^* T^* S} (\phi_{,M}^* L_{,\sigma} - \text{div}_M (\phi_{,M}^* L_{,v})) dV_g + \int_{\partial M} A \cdot_{\phi^* T^* S} \phi_{,M}^* L_{,v} \cdot T^* M \nu d\bar{V}_g.$$

The expression above is often called the **first variation** of  $\mathcal{L}$ . A type analysis here gives  $\phi_{,M}^* L_{,\sigma} \in \Gamma(\phi^* T^* S)$  and  $\phi_{,M}^* L_{,v} \in \Gamma(\phi^* T^* S \otimes_M TM)$ . Recall that because the domain of  $\phi$  is  $M$ ,  $\nabla \phi \equiv \phi_{,M}$ .

*Proof.* Supporting calculations will be made below in lemmas. Let  $\Phi: M \times I \rightarrow S$  be as in (11.1), so that  $\delta_i \Phi = A$ . For tidiness, let  $\mathbf{L}_{,\sigma} := \phi_{,M}^* L_{,\sigma}$  and  $\mathbf{L}_{,v} := \phi_{,M}^* L_{,v}$ . Then

$$\begin{aligned}
d\mathcal{L}(\phi) \cdot A &= d\mathcal{L}(\phi) \cdot \delta_i \Phi \\
&= \delta_i(\mathcal{L}(\Phi)) \\
&= \int_M \delta_i(L \circ \Phi_{,M}) dV_g \\
&= \int_M \mathbf{L}_{,\sigma} \cdot \phi^* T^* S A + \mathbf{L}_{,v} \cdot \phi^* T^* S \otimes_M T^* M \nabla^{\phi^* T^* S} A dV_g && \text{(by (12.2))} \\
&= \int_M A \cdot \phi^* T^* S (\mathbf{L}_{,\sigma} - \text{div}_M \mathbf{L}_{,v}) + \text{div}_M (A \cdot \phi^* T^* S \mathbf{L}_{,v}) dV_g && \text{(by (12.2))} \\
&= \int_M A \cdot \phi^* T^* S (\mathbf{L}_{,\sigma} - \text{div}_M \mathbf{L}_{,v}) dV_g + \int_{\partial M} A \cdot \phi^* T^* S \mathbf{L}_{,v} \cdot T^* M \nu d\bar{V}_g \quad \text{(divergence theorem),}
\end{aligned}$$

as desired.

As for the types of  $\phi_{,M}^* L_{,\sigma}$  and  $\phi_{,M}^* L_{,v}$ , the contravariance of bundle pullback allows significant simplification. Because  $L_{,\sigma} \in \Gamma(\pi_S^* T^* S)$  and  $L_{,v} \in \Gamma(\pi^* E^*)$ ,

$$\begin{aligned}
\phi_{,M}^* L_{,\sigma} &\in \Gamma(\phi_{,M}^* \pi_S^* T^* S) = \Gamma((\pi_S \circ \phi_{,M})^* T^* S) = \Gamma(\phi^* T^* S) \quad \text{and} \\
\phi_{,M}^* L_{,v} &\in \Gamma(\phi_{,M}^* \pi^* E^*) = \Gamma((\pi \circ \phi_{,M})^* (T^* S \otimes TM)) = \Gamma(\phi^* T^* S \otimes_M TM).
\end{aligned}$$

The supporting calculations follow. Define  $z: M \rightarrow M \times I$ ,  $m \mapsto (m, 0)$  for purposes of evaluation of  $i = 0$  via precomposition as in (10.6). Then  $\delta_i$  is a section of a pullback bundle;  $\delta_i = z^* \partial_i \in \Gamma(z^* (TM \oplus TI))$ . It should be noted that  $\Phi \circ z = \phi$  by definition, and that  $z^* \Phi_{,M} = (z^* \Phi)_{,M} = \phi_{,M}$  by (10.6).  $\square$

**Lemma 12.2.** *Let  $L$ ,  $\Phi$ ,  $A$ ,  $\sigma$ , and  $v$  be as in Theorem 12.1. The variational derivative of  $L \circ \Phi_{,M}$  decomposes in terms of the partial covariant derivatives  $\mathbf{L}_{,\sigma}$  and  $\mathbf{L}_{,v}$  and the linearized variation  $A$ ;*

$$\delta_i(L \circ \Phi_{,M}) = \phi_{,M}^* L_{,\sigma} \cdot \phi^* T^* S \delta_i \Phi + \phi_{,M}^* L_{,v} \cdot (\phi \times_M \text{Id}_M)^* E \nabla^\phi \delta_i \Phi.$$

The integration-by-parts trick as in the derivation of the first variation in elementary calculus of variations generalizes to the covariant setting;

$$\mathbf{L}_{,\sigma} \cdot \phi^* T^* S A + \mathbf{L}_{,v} \cdot \phi^* T^* S \otimes_M T^* M \nabla^\phi A = A \cdot \phi^* T^* S (\mathbf{L}_{,\sigma} - \text{div}_M \mathbf{L}_{,v}) + \text{div}_M (A \cdot \phi^* T^* S \mathbf{L}_{,v}).$$

*Proof.* A wonderful string of equalities follows.

$$\begin{aligned}
&\delta_i(L \circ \Phi_{,M}) \\
&= z^* \nabla^{M \times I \rightarrow \mathbb{R}} (L \circ \Phi_{,M}) \cdot z^* (TM \oplus TI) \delta_i && \text{(here, } \delta_i = z^* \partial_i) \\
&= z^* \Phi_{,M}^* \nabla^{E \rightarrow \mathbb{R}} L \cdot z^* \Phi_{,M}^* T^* E z^* \nabla^{M \times I \rightarrow E} \Phi_{,M} \cdot z^* (TM \oplus TI) \delta_i && \text{(chain rule)} \\
&= \phi_{,M}^* (L_{,\sigma} \cdot \pi_S^* T^* S \sigma + L_{,\mu} \cdot \pi_M^* T^* M \mu + L_{,v} \cdot \pi^* E v) \cdot \phi_{,M}^* T^* E \delta_i \Phi_{,M} \quad \text{(by (9.4) and because } \Phi_{,M} \circ z = \phi_{,M}) \\
&= \phi_{,M}^* L_{,\sigma} \cdot \phi_{,M}^* \pi_S^* T^* S \phi_{,M}^* \sigma \cdot \phi_{,M}^* T^* E \delta_i \Phi_{,M} \\
&\quad + \phi_{,M}^* L_{,\mu} \cdot \phi_{,M}^* \pi_M^* T^* M \phi_{,M}^* \mu \cdot \phi_{,M}^* T^* E \delta_i \Phi_{,M} \\
&\quad + \phi_{,M}^* L_{,v} \cdot \phi_{,M}^* \pi^* E \phi_{,M}^* v \cdot \phi_{,M}^* T^* E \delta_i \Phi_{,M} \\
&= \phi_{,M}^* L_{,\sigma} \cdot \phi^* T^* S \delta_i \Phi + \phi_{,M}^* L_{,v} \cdot (\phi \times_M \text{Id}_M)^* E \nabla^\phi \delta_i \Phi && \text{(by (12.3))}
\end{aligned}$$

Note that by (6.3),  $\Phi_{*,M}^* \pi_S^* TS = (\pi_S \circ \Phi_{*,M})^* TS = \Phi^* TS$ ,  $\Phi_{*,M}^* \pi_M^* TM = (\pi_M \circ \Phi_{*,M})^* TM = (\text{pr}_M^{M \times I})^* TM$  and  $\Phi_{*,M}^* \pi^* E = (\pi \circ \Phi_{*,M})^* E = (\Phi \times_{M \times I} \text{pr}_M^{M \times I})^* E$ . Replacing  $\delta_i \Phi$  with  $A$  gives

$$\delta_i (L \circ \Phi_{*,M}) = \mathbf{L}_{,\sigma} \cdot \phi^* TS A + \mathbf{L}_{,v} \cdot \phi^* TS \otimes_M T^* M \nabla^\phi A,$$

establishing the first equality.

For the second,

$$\begin{aligned} & \mathbf{L}_{,\sigma} \cdot \phi^* TS A + \mathbf{L}_{,v} \cdot \phi^* TS \otimes_M T^* M \nabla^\phi A \\ &= \mathbf{L}_{,\sigma} \cdot \phi^* TS A + \text{tr}_{T^* M} (\mathbf{L}_{,v} \cdot \phi^* TS \nabla^\phi A) && \text{(tracing } TM \text{ separately)} \\ &= A \cdot \phi^* T^* S \mathbf{L}_{,\sigma} + \text{tr}_{T^* M} (\nabla (\mathbf{L}_{,v} \cdot \phi^* TS A) - (\nabla^{\phi \times_M \text{Id}_M} \mathbf{L}_{,v}) \cdot \phi^* TS A) && \text{(reverse product rule)} \\ &= A \cdot \phi^* T^* S \mathbf{L}_{,\sigma} - A \cdot \phi^* T^* S \text{tr}_{T^* M} \nabla^{\phi \times_M \text{Id}_M} \mathbf{L}_{,v} + \text{tr}_{T^* M} \nabla (A \cdot \phi^* T^* S \mathbf{L}_{,v}) && (\cdot \phi^* TS \text{ commutes with } \text{tr}_{T^* M}) \\ &= A \cdot \phi^* T^* S (\mathbf{L}_{,\sigma} - \text{div}_M \mathbf{L}_{,v}) + \text{div}_M (A \cdot \phi^* T^* S \mathbf{L}_{,v}) && \text{(definition of } \text{div}_M \text{)}. \end{aligned}$$

Note that  $\mathbf{L}_{,v} \in \Gamma(\phi^* T^* S \otimes_M TM)$ , so  $\text{div}_M \mathbf{L}_{,v} \in \Gamma(\phi^* T^* S)$  and  $A \cdot \phi^* T^* S \mathbf{L}_{,v} \in \Gamma(TM)$ .  $\square$

**Lemma 12.3.** *The variation  $\delta_i \Phi_{*,M}$  decomposes as follows.*

$$\begin{aligned} \phi_{*,M}^* \sigma \cdot \phi_{*,M}^* TE \delta_i \Phi_{*,M} &= \delta_i \Phi \in \Gamma(\phi^* TS), \\ \phi_{*,M}^* \mu \cdot \phi_{*,M}^* TE \delta_i \Phi_{*,M} &= 0 \in \Gamma(TM), \\ \phi_{*,M}^* v \cdot \phi_{*,M}^* TE \delta_i \Phi_{*,M} &= \nabla^{\phi^* TS} \delta_i \Phi \in \Gamma(\phi^* TS \otimes_M T^* M). \end{aligned}$$

*Proof.* This calculation determines the  $\sigma$  component of  $\delta_i \Phi_{*,M}$ .

$$\begin{aligned} \phi_{*,M}^* \sigma \cdot \phi_{*,M}^* TE \delta_i \Phi_{*,M} &= \phi_{*,M}^* \nabla \pi_S \cdot \phi_{*,M}^* TE \delta_i \Phi_{*,M} \\ &= \delta_i (\pi_S \circ \Phi_{*,M}) \\ &= \delta_i (\text{pr}_S^{S \times M} \circ \pi \circ \Phi_{*,M}) \\ &= \delta_i \Phi \in \Gamma(z^* \Phi^* TS) \cong \Gamma(\phi^* TS). \end{aligned}$$

This calculation determines the  $\mu$  component of  $\delta_i \Phi_{*,M}$ .

$$\begin{aligned} \phi_{*,M}^* \mu \cdot \phi_{*,M}^* TE \delta_i \Phi_{*,M} &= \phi_{*,M}^* \nabla \pi_M \cdot \phi_{*,M}^* TE \delta_i \Phi_{*,M} \\ &= \delta_i (\pi_M \circ \Phi_{*,M}) \\ &= \delta_i (\text{pr}_M^{S \times M} \circ \pi \circ \Phi_{*,M}) \\ &= \delta_i \text{pr}_M^{M \times I} \\ &= 0 \in \Gamma(z^* (\text{pr}_M^{M \times I})^* TM) \cong \Gamma(TM). \end{aligned}$$

The last equality follows from the fact that  $\text{pr}_M^{M \times I}$  does not depend on the  $i$  coordinate.

This calculation determines the  $v$  component of  $\delta_i \Phi_{*,M}$ . Let  $p_M := \text{pr}_M^{M \times I}$  and  $p_I := \text{pr}_I^{M \times I}$ . The left-hand side of the third equality claimed in the lemma will be examined before evaluating at  $i = 0$ ;

$$\Phi_{*,M}^* v \cdot \Phi_{*,M}^* TE \partial_i \Phi_{*,M} = \nabla_{p_I^* \partial_i}^{(\pi \circ \Phi_{*,M})^* E} \Phi_{*,M} = \nabla_{p_I^* \partial_i}^{(\Phi \times_{M \times I} p_M)^* (TS \otimes T^* M)} \Phi_{*,M} \in \Gamma(\Phi^* TS \otimes_{M \times I} p_M^* T^* M).$$



Let  $Y \in \Gamma(TM)$ , noting that  $p_M^* Y \in \Gamma(p_M^* TM)$ . Then

$$\begin{aligned}
& \left( \Phi_{,M}^* v \cdot \Phi_{,M}^* TE \partial_i \Phi_{,M} \right) \cdot p_M^* TM p_M^* Y \\
&= \nabla_{p_I^* \partial_i}^{\Phi^* TS \otimes M \times I p_M^* T^* M} \Phi_{,M} \cdot p_M^* TM p_M^* Y \\
&= \left( \Phi_{,MI} \cdot p_I^* TI p_I^* \partial_i \right) \cdot p_M^* TM p_M^* Y \\
&= \left( \Phi_{,IM} \cdot p_M^* TM p_M^* Y \right) \cdot p_I^* TI p_I^* \partial_i \quad (\text{by (10.3)}) \\
&= \left( \Phi_{,I} \cdot p_I^* TI p_I^* \partial_i \right)_{,M} \cdot p_M^* TM p_M^* Y - \Phi_{,I} \cdot p_I^* TI \left( (p_I^* \partial_i)_{,M} \cdot p_M^* TM p_M^* Y \right) \\
&= (\partial_i \Phi)_{,M} \cdot p_M^* TM p_M^* Y \quad (\text{since } p_I^* \partial_i \text{ doesn't depend on } M).
\end{aligned}$$

Recall that  $\text{Id}_M = p_M \circ z$  and that the pullback of bundles is contravariant. Then evaluating at  $i = 0$  via pullback by  $z$  renders

$$\begin{aligned}
& \left( \phi_{,M}^* v \cdot \phi_{,M}^* TE \delta_i \Phi_{,M} \right) \cdot TM Y \\
&= \left( (\Phi_{,M} \circ z)^* v \cdot (\Phi_{,M} \circ z)^* TE z^* \partial_i \Phi_{,M} \right) \cdot (p_M \circ z)^* TM (p_M \circ z)^* Y \\
&= \left( z^* \Phi_{,M}^* v \cdot z^* \Phi_{,M}^* TE z^* \partial_i \Phi_{,M} \right) \cdot z^* p_M^* TM z^* p_M^* Y \\
&= z^* \left( \left( \Phi_{,M}^* v \cdot \Phi_{,M}^* TE \partial_i \Phi_{,M} \right) \cdot p_M^* TM p_M^* Y \right) \\
&= z^* \left( (\partial_i \Phi)_{,M} \cdot p_M^* TM p_M^* Y \right) \\
&= z^* (\partial_i \Phi)_{,M} \cdot z^* p_M^* TM z^* p_M^* Y \\
&= (z^* \partial_i \Phi)_{,M} \cdot (p_M \circ z)^* TM (p_M \circ z)^* Y \quad (\text{by (10.6)}) \\
&= (\delta_i \Phi)_{,M} \cdot TM Y \\
&= \nabla^{\phi^* TS} \delta_i \Phi \cdot TM Y
\end{aligned}$$

The last equality is because  $\delta_i \Phi \in \Gamma(\phi^* TS)$ , which is a bundle over  $M$ , and therefore  $(\delta_i \Phi)_{,M}$  is the total covariant derivative. Because  $Y$  is pointwise-arbitrary in  $TM$ , this implies that  $\phi_{,M}^* \cdot \phi_{,M}^* TI \delta_i \Phi_{,M} = \nabla^{\phi^* TS} \delta_i \Phi$ , i.e. the variational derivative  $\delta_i$  commutes with the first material derivative, just as in the analogous situation in elementary calculus of variations.  $\square$

**Corollary 12.4** (Euler-Lagrange equations). *If  $\phi \in C^\infty(M, S)$  is a critical point of  $\mathcal{L}$  (i.e. if  $d\mathcal{L}(\phi) \cdot A = 0$  for all  $A \in \Gamma(\phi^* TS)$ ), then*

$$\begin{aligned}
\phi_{,M}^* L_{,\sigma} - \text{div}_M (\phi_{,M}^* L_{,v}) &= 0 \text{ on } M, \\
\phi_{,M}^* L_{,v} \cdot TM \nu &= 0 \text{ on } \partial M.
\end{aligned}$$

These are called the **Euler-Lagrange equations** for the energy functional  $\mathcal{L}$ . Recall that because the domain of  $\phi$  is  $M$ ,  $\nabla \phi \equiv \phi_{,M}$ .

*Proof.* This follows trivially from (12.1) and the Fundamental Lemma of the Calculus of Variations [6, pg. 16].  $\square$

It should be noted that the boundary Euler-Lagrange equation is due to the fact that the admissible variations are entirely unrestricted. If, for example, the class of maps being considered had fixed boundary data, then any variation would vanish at the boundary, and there would be no boundary Euler-Lagrange equation; this is typically how geodesics and harmonic maps are formulated.

*Remark 12.5* (Analogues in elementary calculus of variations). The quantities  $L_{,\mu}, L_{,\sigma}, L_{,v}$  generalize the quantities  $\frac{\partial L}{\partial x}, \frac{\partial L}{\partial z}, \frac{\partial L}{\partial p}$  respectively of the elementary treatment of the calculus of variations for energy functional

$$(f: U \rightarrow \mathbb{R}^n) \mapsto \int_U L(x, f(x), Df(x)) dx,$$

where  $U \subset \mathbb{R}^m$  is compact and  $U \times \mathbb{R}^n \times \mathbb{R}^{m \times n} \ni (x, z, p) \mapsto L(x, z, p)$  is the Lagrangian. Here,  $\frac{\partial L}{\partial x}: U \times \mathbb{R}^n \times \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^m$ ,  $\frac{\partial L}{\partial z}: U \times \mathbb{R}^n \times \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^n$ , and  $\frac{\partial L}{\partial p}: U \times \mathbb{R}^n \times \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$  decompose the total derivative  $dL$  and are defined by the relation

$$dL(x, z, p) \cdot (u, v, w) = \frac{\partial L}{\partial x}(x, z, p) \cdot u + \frac{\partial L}{\partial z}(x, z, p) \cdot v + \frac{\partial L}{\partial p}(x, z, p) : w$$

for  $u \in \mathbb{R}^m$ ,  $v \in \mathbb{R}^n$ , and  $w \in \mathbb{R}^{m \times n}$ . The Euler-Lagrange equation in this setting is

$$\left( \frac{\partial L}{\partial z} - \operatorname{div}_U \frac{\partial L}{\partial p} \right) (x, f(x), Df(x)) = 0 \text{ for } x \in U,$$

noting that the left hand side of the equation takes values in  $\mathbb{R}^n$ .

In most situations involving simpler calculations, it is desirable and acceptable to dispense with the highly decorated notation and use trimmed-town, context-dependent notation, leaving off type-specifying sub/superscripts when clear from context.

**Proposition 12.6** (Conserved quantity). *If  $M$  is a real interval,  $\phi \in C^\infty(M, S)$  satisfies the Euler-Lagrange equation, and  $L_{,\mu} = 0$ , then*

$$H := (\nabla \phi)^* L_{,v} \cdot \phi^* TS \otimes_M T^* M \nabla \phi - (\nabla \phi)^* L \in C^\infty(M, \mathbb{R})$$

*is constant. If  $L$  is kinetic minus potential energy, then  $H$  is kinetic plus potential energy (the total energy), and is referred to as the Hamiltonian.*

*Proof.* Let  $t$  be the standard real coordinate. Note that because  $M$  is a real interval,  $\nabla \phi = \phi' \otimes_M dt$ . Terms appearing in the derivative of  $H$  can be simplified as follows. Note the repeated  $\nabla$  derivatives;  $\nabla \phi: M \rightarrow \phi^* TS \otimes_M T^* M$  but  $\nabla \nabla \phi: M \rightarrow (\nabla \phi)^* T(\phi^* TS \otimes_M T^* M)$  lands in a higher tangent space.

$$\begin{aligned} (\nabla \phi)^* \sigma \cdot \nabla \nabla \phi \cdot \frac{d}{dt} &= (\nabla \phi)^* \nabla \pi_S \cdot \nabla \nabla \phi \cdot \frac{d}{dt} = \frac{d}{dt} (\pi_S \circ \nabla \phi) = \phi', \\ (\nabla \phi)^* v \cdot \nabla \nabla \phi \cdot \frac{d}{dt} &= (\nabla \phi)^* v \cdot \frac{d}{dt} \nabla \phi = \nabla_{\frac{d}{dt}} \nabla \phi, \\ \nabla_{\frac{d}{dt}} (\nabla \phi)^* L &= (\nabla \phi)^* \nabla L \cdot \nabla \nabla \phi \cdot \frac{d}{dt} \\ &= (\nabla \phi)^* L_{,\sigma} \cdot \phi' + (\nabla \phi)^* L_{,v} \cdot \nabla_{\frac{d}{dt}} \nabla \phi, \\ \nabla_{\frac{d}{dt}} ((\nabla \phi)^* L_{,v} : \nabla \phi) &= \nabla_{\frac{d}{dt}} (\nabla \phi)^* L_{,v} : (\phi' \otimes_M dt) + (\nabla \phi)^* L_{,v} : \nabla_{\frac{d}{dt}} \nabla \phi \\ &= \left( \nabla_{\frac{d}{dt}} (\nabla \phi)^* L_{,v} \cdot dt \right) \cdot \phi' + (\nabla \phi)^* L_{,v} : \nabla_{\frac{d}{dt}} \nabla \phi. \end{aligned}$$

Again, because  $M$  is a real interval, the divergence is just the derivative, so the Euler-Lagrange equation is

$$\begin{aligned} 0 &= (\nabla \phi)^* L_{,\sigma} - \operatorname{div}_M ((\nabla \phi)^* L_{,v}) \\ &= (\nabla \phi)^* L_{,\sigma} - \nabla (\nabla \phi)^* L_{,v} : \left( dt \otimes_M \frac{d}{dt} \right) \\ &= (\nabla \phi)^* L_{,\sigma} - \nabla_{\frac{d}{dt}} (\nabla \phi)^* L_{,v} \cdot dt, \end{aligned}$$

and therefore  $\nabla_{\frac{d}{dt}} (\nabla \phi)^* L_{,v} \cdot dt = (\nabla \phi)^* L_{,\sigma}$ . Thus

$$\nabla_{\frac{d}{dt}} H = \nabla_{\frac{d}{dt}} ((\nabla \phi)^* L_{,v} : \nabla \phi - (\nabla \phi)^* L) = \left( \nabla_{\frac{d}{dt}} (\nabla \phi)^* L_{,v} \cdot dt - (\nabla \phi)^* L_{,\sigma} \right) \cdot \phi'$$

which is zero because  $\phi$  satisfies the Euler-Lagrange equation. This shows that  $H$  is constant along solutions of the Euler-Lagrange equation, and is therefore a conserved quantity. It should be noted that this proof relies on the fact that the divergence takes a particularly simple form when the domain  $M$  is a real interval; the result does not necessarily hold for a general choice of  $M$ .  $\square$

**Example 12.7** (Harmonic maps). Define a metric

$$k \in \Gamma(E^* \otimes_{S \times M} E^*) \cong \Gamma((TS \otimes T^*M) \otimes_{S \times M} (TS \otimes T^*M))$$

in a manner analogous to that in (3.4);

$$k := h \boxtimes g^{-1}.$$

To clarify,  $h \otimes g^{-1} \in \Gamma((T^*S \otimes_S T^*S) \otimes (TM \otimes_M TM))$ , so permuting the middle two components (as in the definition of  $h \boxtimes g^{-1}$ ) gives the correct type, including the necessary metric symmetry condition. If  $A \in E$ , then  $|A|_k^2$  is the quantity obtained by raising/lowering the indices of  $A$  and pairing it naturally with  $A$ . A useful fact is that  $\nabla k = 0$ ; if  $u \oplus v \in TS \oplus TM$ , then permutation commutativity (8.12) and the product rule gives

$$\nabla_{u \oplus v} k = \nabla_{u \oplus v} (h \boxtimes g^{-1}) = \nabla_u h \boxtimes g^{-1} + h \boxtimes \nabla_v g^{-1},$$

which equals zero because  $h$  and  $g^{-1}$  are parallel with respect to  $\nabla^{TS}$  and  $\nabla^{T^*M}$  respectively.

With Lagrangian

$$L : E \rightarrow \mathbb{R}, A \mapsto \frac{1}{2} |A|_k^2$$

and energy functional

$$\mathcal{E}(\phi) := \int_M L \circ \nabla \phi dV_g$$

( $\mathcal{E}(\phi)$  is called the **energy** of  $\phi$ ), the resulting Euler-Lagrange equations can be written down after calculating  $L_{,\sigma}$  and  $L_{,v}$ . It is worthwhile to note that  $L$  is a quadratic form  $A \mapsto A : \frac{1}{2}k : A$  on  $E$ , which will automatically imply that  $L_{,v}(A) = A : k$ . However, the calculation showing this will be carried out for demonstration purposes.

Let  $A, B \in TS \otimes T^*M$ . Then  $\epsilon \mapsto A + \epsilon B$  is a vertical variation of  $A$ , since  $h(\delta(A + \epsilon B)) = 0$ , so

$$\begin{aligned} L_{,v}(A) : B &= L_{,v}(A) : v \cdot \delta_\epsilon(A + \epsilon B) \\ &= \delta_\epsilon(L(A + \epsilon B)) \\ &= \delta_\epsilon \left( (A + \epsilon B) : \frac{1}{2}(\pi^*k(A + \epsilon B)) : (A + \epsilon B) \right). \end{aligned}$$

The product rule gives three terms. The middle term is zero because  $\pi(A + \epsilon B) = \pi(A)$ , and therefore does not depend on  $\epsilon$ . The basepoint evaluation notation for  $\pi^*k(A)$  will be suppressed for brevity (see Section 5). Thus

$$L_{,v}(A) : B = B : \frac{1}{2}k : A + A : \frac{1}{2}k : B = A : k : B,$$

where the last equality results from the symmetry of  $k$ . By the nondegeneracy of the natural pairing on  $TS \otimes T^*M$  (which is denoted here by  $:$ ), this implies that  $L_{,v}(A) = A : k$ .

To calculate  $L_{,\sigma}$ , it is sufficient (and can be easier) to calculate  $L_{,h}$ , as  $h = \nabla \pi$ ,  $\pi = \pi_S \times_E \pi_M$ , so  $h = \sigma \oplus_E \mu$ . Let  $A(\epsilon)$  be a horizontal curve in  $E = TS \otimes T^*M$ ; this means that  $v \cdot \frac{d}{d\epsilon} A = 0$ . Recall that

$v \cdot \frac{d}{d\epsilon} A$  is defined by  $\nabla^{(\pi \circ A)^* E} A$ . Then

$$\begin{aligned} L_{,h}(A) \cdot \pi^*(TS \oplus TM) h \cdot_{TE} \delta_\epsilon A &= (L_{,h} \cdot \pi^*(TS \oplus TM) h + L_{,v} \cdot \pi^* E v) \cdot_{TE} \delta_\epsilon A \\ &= \nabla L \cdot_{TE} \delta_\epsilon A \\ &= \delta_\epsilon (L \circ A) \\ &= \delta_\epsilon \left( A : \frac{1}{2} A^* \pi^* k : A \right). \end{aligned}$$

As before, the product rule gives three terms. Using the contravariance of bundle pullback, the middle term is

$$\frac{1}{2} \nabla_{\delta_\epsilon}^{(\pi \circ A)^* (E^* \otimes_{S \times M} E^*)} (\pi \circ A)^* k = \frac{1}{2} (\pi \circ A)^* \nabla^{E^* \otimes_{S \times M} E^*} k \cdot \delta_\epsilon (\pi \circ A),$$

which equals zero because  $\nabla k = 0$ . Thus

$$L_{,h}(A) \cdot h \cdot \delta_\epsilon A = \nabla_{\delta_\epsilon} A : \frac{1}{2} k : A + A : \frac{1}{2} k : \nabla_{\delta_\epsilon} A,$$

which equals zero because  $\nabla_{\delta_\epsilon} A = v \cdot \delta_\epsilon A = 0$ . The quantity  $h \cdot \delta_\epsilon A$  can take any value in  $\pi^*(TS \oplus TM)$ , showing that  $L_{,h} = 0$ . Finally,  $h = \sigma \oplus_E \mu$  implies that  $L_{,\sigma} = 0$  and  $L_{,\mu} = 0$ . This can be understood from the fact that  $L$  depends only on the fiber values of  $A$ , and has no explicit dependence on the basepoint; this relies crucially on the fact that  $\nabla k = 0$ .

Finally, the Euler-Lagrange equations can be written down. Recalling that the natural trace of a tensor (used in the divergence term in the Euler-Lagrange equation) is contraction with the appropriate identity tensor, let  $(e_i)$  be a local frame for  $TM$  and let  $(e^i)$  be its dual coframe, so that  $e_i \otimes_M e^i$  is a local expression<sup>5</sup> for  $\text{Id}_{TM} \in \Gamma(TM \otimes_M T^*M)$ . The type-subscripted notation will be minimized except to help clarify. On  $M$ :

$$\begin{aligned} 0 &= (\nabla \phi)^* L_{,\sigma} - \text{div}_M ((\nabla \phi)^* L_{,v}) \\ &= -\text{tr} \nabla (\nabla \phi : k) \\ &= -\nabla_{e_i} (\nabla \phi : k) \cdot_{T^*M} e^i \\ &= -\nabla_{e_i} \nabla \phi : k \cdot e^i - \nabla \phi : \nabla_{e_i} k \cdot e^i. \end{aligned}$$

The second term vanishes because  $\nabla k = 0$ . Unraveling the definition of  $k$  gives  $\nabla_{e_i} \nabla \phi : k = \phi^* h \cdot \nabla_{e_i} \nabla \phi \cdot g^{-1}$ . Contracting both sides of the above equation with  $-\phi^* h^{-1}$  gives

$$0 = \nabla_{e_i} \nabla \phi \cdot g^{-1} \cdot e^i = \nabla_{e_i} \nabla \phi \cdot e_i = \text{tr}_g \nabla^2 \phi \in \Gamma(\phi^* TS).$$

The quantity  $\text{tr}_g \nabla^2 \phi$  is the  $g$ -trace of the covariant Hessian of  $\phi$  and can rightfully be called the **covariant Laplacian** of  $\phi$  and denoted by  $\Delta_g \phi$  (this is also referred to as the **tension field** of  $\phi$  in other literature [20, pg. 13], which is denoted  $\tau(\phi)$ ). Note that  $\Delta_g \phi$  is a vector field along  $\phi$ . This makes sense because  $\phi$  is not necessarily a scalar function; it takes values in  $S$ . In the case  $S = \mathbb{R}$ ,  $\Delta_g \phi$  is the ordinary covariant Laplacian on scalar functions.

A **harmonic map** is defined as a critical point of the energy functional  $\mathcal{E}(\phi) := \int_M \frac{1}{2} |\nabla \phi|_k^2 dV_M$ . Assuming a fixed boundary (so that the variations vanish on the boundary) eliminates the boundary Euler-Lagrange equation, the remaining equation is

$$\Delta_g \phi = 0 \text{ on the interior of } M,$$

which is the generalization of Laplace's equation. Satisfying Laplace's equation is a sufficient condition for a map to be a critical point of the energy functional. There is an abundance of literature concerning harmonic maps and the analysis thereof [2, 6, 14, 20].

<sup>5</sup>It should be noted that while  $\text{Id}_{TM}$  is being written as the local expression  $e_i \otimes_M e^i$ , no inherently local property is being used; this tensor decomposition is only used so that the product rule can be used in the following calculations in a clear way.

**Example 12.8** (The geodesic equation). A fundamental problem in differential geometry is determining length-minimizing curves between given points. If  $M$  is a bounded, real interval, and  $t$  denotes the standard real coordinate, then the length functional on curves  $\phi: M \rightarrow S$  is  $\mathcal{L}(\phi) := \int_M |\phi'|_g dt$ . A topological metric  $d: M \times M \rightarrow \mathbb{R}$  on  $M$  can be defined as

$$d(p, q) := \inf \{ \mathcal{L}(\phi) \mid \phi \text{ joins } p \text{ to } q \}.$$

It can be shown that the length functional  $\mathcal{L}(\phi) := \int_M |\phi'|_h dt$  and the energy functional  $\mathcal{E}(\phi) := \int_M \frac{1}{2} |\phi'|_h^2 dt$  have identical minimizers. Note that  $\phi' \in \Gamma(\phi^*TS)$ . It is therefore sufficient to consider the analytically preferable energy functional.

In this case, the metric  $g$  on  $M$  is just scalar multiplication on  $\mathbb{R}$ . Because  $M$  is one-dimensional and  $t$  is the standard real coordinate,  $\frac{d}{dt}$  is a global, parallel orthonormal frame for  $TM$ , and the  $g$ -trace of  $\nabla^2\phi$  (i.e.  $\Delta_g\phi$ ) has a single term. The Euler-Lagrange equation, on the interior of  $M$ , is

$$0 = \Delta_g\phi = \text{tr}_g \nabla^2\phi = \nabla \nabla \phi : \left( \frac{d}{dt}, \frac{d}{dt} \right) = \nabla_{\frac{d}{dt}} \nabla \phi \cdot \frac{d}{dt} = \nabla_{\frac{d}{dt}} \left( \nabla \phi \cdot \frac{d}{dt} \right) - \nabla \phi \cdot \nabla_{\frac{d}{dt}} \frac{d}{dt}.$$

But  $\nabla \phi \cdot \frac{d}{dt} = \phi'$  and  $\nabla_{\frac{d}{dt}} \frac{d}{dt} = 0$ , giving the **geodesic equation**

$$\nabla_{\frac{d}{dt}}^{\phi^*TS} \phi' = 0 \text{ on the interior of } M.$$

This is the covariant way to state that the acceleration of  $\phi$  is identically zero. The geodesic equation is commonly notated as  $0 = \nabla_{\phi'} \phi'$ , though such notation is inaccurate because  $\phi'$  is not a vector field on  $S$ , but a vector field along  $\phi$ , and therefore use of the pullback covariant derivative  $\nabla^{\phi^*TS}$  is correct (see (8.8)).

While formulated using fixed boundary conditions ( $\phi$  has  $p$  and  $q$  as its endpoints), the geodesic equation is a second order ODE for which initial tangent vector conditions are sufficient to uniquely determine a solution.

## 13 Second Variation

A further consideration after finding critical points of the energy functional  $\mathcal{L}$  is determining which critical points are extrema. This will involve calculating the second derivative of  $\mathcal{L}$ . Let  $C := C^\infty(M, S)$ , noting that  $T_\phi C \cong \Gamma(\phi^*TS)$  for  $\phi \in C$ . The first derivative of  $\mathcal{L}$  is  $\nabla^{C \rightarrow \mathbb{R}} \mathcal{L} := d\mathcal{L}$ , as seen in the previous section. The second derivative is the covariant Hessian  $\nabla^{T^*C} \nabla^{C \rightarrow \mathbb{R}} \mathcal{L}$ , where the covariant derivative  $\nabla^{T^*C}$  is induced by  $\nabla^{TS}$  [5, Theorem 5.4].

For the remainder of this section, let  $I, J \subseteq \mathbb{R}$  be neighborhoods of zero, let  $i$  and  $j$  be their respective standard coordinates, and extend the existing  $\delta$ -style derivative-at-a-point notation by defining  $\delta_i := \frac{\partial}{\partial i} |_{i=j=0}$ ,  $\delta_j := \frac{\partial}{\partial j} |_{i=j=0}$ , and evaluation map  $z: M \rightarrow M \times I \times J$ ,  $m \mapsto (m, 0, 0)$ . Then  $\delta_i = z^* \partial_i$  and  $\delta_j = z^* \partial_j$ ; these will be used as in the calculation of the first variation.

**Theorem 13.1** (Second variation of  $\mathcal{L}$ ). *Let  $\mathcal{L}$ ,  $L$ ,  $\sigma$ ,  $\mu$ ,  $v$  and  $\nu$  all be defined as above. If  $\phi \in C^\infty(M, S)$  is a critical point of  $\mathcal{L}$  and  $A, B \in T_\phi C \cong \Gamma(\phi^*TS)$ , then the covariant Hessian of  $\mathcal{L}$  is*

$$\begin{aligned} & \nabla^2 \mathcal{L}(\phi) :_{T_\phi C} (A \otimes B) \\ &= \int_M A \cdot \phi^* T^* S \phi_{,M}^* L_{,\sigma\sigma} \cdot \phi^* TS B + A \cdot \phi^* T^* S \phi_{,M}^* L_{,\sigma\nu} \cdot \phi^* TS \otimes_M T^* M \nabla^{\phi^*TS} B \\ & \quad + \nabla^{\phi^*TS} A \cdot \phi^* T^* S \otimes_M TM \phi_{,M}^* L_{,\nu\sigma} \cdot \phi^* TS B + \nabla^{\phi^*TS} A \cdot \phi^* T^* S \otimes_M TM \phi_{,M}^* L_{,\nu\nu} \cdot \phi^* TS \otimes_M T^* M \nabla^{\phi^*TS} B \\ & \quad - A \cdot \phi^* T^* S (\phi_{,M}^* L_{,\nu} \cdot \phi^* TS \otimes_M T^* M (\phi^* R^{TS} \cdot \phi^* TS \phi_{,M})) \cdot \phi^* TS B dV_g. \end{aligned}$$

This is often called the **second variation** of  $\mathcal{L}$ . Here,  $R^{TS} \in \Gamma(TS \otimes_S T^*S \otimes_S T^*S \otimes_S T^*S)$  denotes the Riemannian curvature endomorphism tensor for the Levi-Civita connection on  $TS$ .

*Proof.* Let  $\Phi: M \times I \times J \rightarrow S$  be a two-parameter variation such that  $\delta_i \Phi = A$  and  $\delta_j \Phi = B$  (e.g.  $\Phi(m, i, j) := \exp(iA(m) + jB(m))$ ). The variation  $\Phi$  can be naturally identified with a variation  $\bar{\Phi}: I \times J \rightarrow C$ ,  $(i, j) \mapsto (m \mapsto \Phi(m, i, j))$  which is more conducive to the use of  $C$  as a manifold. The tensor products in the generally infinite-dimensional  $TC$  are taken formally. Let  $\bar{z} := (0, 0) \in I \times J$ .

By (10.4), taking the algebra formally in the case of infinite-rank vector bundles,

$$\nabla^2(\mathcal{L} \circ \bar{\Phi}) = \bar{\Phi}^* \nabla^2 \mathcal{L} \cdot_{\bar{\Phi}^* TC} (\nabla \bar{\Phi} \boxtimes_{I \times J} \nabla \bar{\Phi}) + \bar{\Phi}^* \nabla \mathcal{L} \cdot_{\bar{\Phi}^* TC} \nabla \nabla \bar{\Phi},$$

so

$$\begin{aligned} & (\nabla^2 \mathcal{L} \circ_C \phi) :_{T_\phi C} (A \otimes B) \\ &= (\nabla^2 \mathcal{L} \circ_C \phi) :_{T_\phi C} (\delta_i \bar{\Phi} \otimes \delta_j \bar{\Phi}) + (\nabla \mathcal{L} \circ_C \phi) \cdot_{T_\phi C} \nabla \bar{\Phi}^{*TC} \partial_i \bar{\Phi} \quad (\text{since } \nabla \mathcal{L} \circ_C \phi = 0) \\ &= (\nabla^2 \mathcal{L} \circ_C \bar{\Phi} \circ_{I \times J} \bar{z}) :_{\bar{z}^* \bar{\Phi}^* TC} (\delta_i \bar{\Phi} \otimes \delta_j \bar{\Phi}) + (\nabla \mathcal{L} \circ_C \bar{\Phi} \circ_{I \times J} \bar{z}) :_{\bar{z}^* \bar{\Phi}^* TC} \nabla \bar{\Phi}^{*TC} \partial_i \bar{\Phi} \\ &= \nabla^2(\mathcal{L} \circ_C \bar{\Phi}) :_{TI \oplus TJ} (\delta_i \otimes_{I \times J} \delta_j) \quad (\text{by above}) \\ &= \delta_j \partial_i (\mathcal{L} \circ_C \bar{\Phi}) \\ &= \int_M \delta_j \partial_i (L \circ \Phi, M) dV_g \\ &= \int_M \nabla^2 (L \circ \Phi, M) :_{TM \oplus TI \oplus TJ} (\delta_i \otimes_{M \times I \times J} \delta_j) dV_g \\ &= \int_M A \cdot_{\phi^* T^* S} \phi_{,M}^* L_{,\sigma\sigma} \cdot_{\phi^* TS} B + A \cdot_{\phi^* T^* S} \phi_{,M}^* L_{,\sigma v} \cdot_{\phi^* TS \otimes_M T^* M} \nabla^{\phi^* TS} B \\ &\quad + \nabla^{\phi^* TS} A \cdot_{\phi^* T^* S \otimes_M TM} \phi_{,M}^* L_{,v\sigma} \cdot_{\phi^* TS} B \\ &\quad + \nabla^{\phi^* TS} A \cdot_{\phi^* T^* S \otimes_M TM} \phi_{,M}^* L_{,vv} \cdot_{\phi^* TS \otimes_M T^* M} \nabla^{\phi^* TS} B \\ &\quad - A \cdot_{\phi^* T^* S} (\phi_{,M}^* L_{,v} \cdot_{\phi^* TS \otimes_M T^* M} (\phi^* R^{TS} \cdot_{\phi^* TS} \phi_{,M})) \cdot_{\phi^* TS} B dV_g \quad (\text{by Calculation (1)}). \end{aligned}$$

Supporting calculations follow.

Calculation (1): Abbreviate  $\phi_{,M}^* L_{,xy}$  by  $\mathbf{L}_{,xy}$ . By (10.4),

$$\begin{aligned} & \nabla^2 (L \circ \Phi, M) :_{TM \oplus TI \oplus TJ} (\delta_i \otimes_{M \times I \times J} \delta_j) \\ &= \left( \left[ \Phi_{,M}^* \nabla^2 L :_{\Phi_{,M}^* TE} (\nabla \Phi, M \boxtimes_{M \times I \times J} \nabla \Phi, M) + \Phi_{,M}^* \nabla L \cdot_{\Phi_{,M}^* TE} \nabla \nabla \Phi, M \right] \circ z \right) :_{z^*(TM \oplus TI \oplus TJ)} (\delta_i \otimes_M \delta_j) \\ &= z^* \Phi_{,M}^* \nabla^2 L :_{z^* \Phi_{,M}^* TE} (\delta_i \Phi, M \otimes_M \delta_j \Phi, M) + z^* \Phi_{,M}^* \nabla L \cdot_{z^* \Phi_{,M}^* TE} \nabla \bar{\Phi}_{,M}^{*TE} \partial_i \Phi, M \\ &\quad (\text{by Calculation (2)}) \\ &= \mathbf{L}_{,\sigma\sigma} :_{\phi_{,M}^* \pi_S^* TS} (\delta_i \Phi \otimes_M \delta_j \Phi) + \mathbf{L}_{,\sigma v} \cdot_{\phi_{,M}^* \pi_S^* TS \otimes_M \phi_{,M}^* \pi^* E} \left( \delta_i \Phi \otimes_M \nabla^{\phi^* TS} \delta_j \Phi \right) \\ &\quad + \mathbf{L}_{,v\sigma} \cdot_{\phi_{,M}^* \pi^* E \otimes_M \phi_{,M}^* \pi_S^* TS} \left( \nabla^{\phi^* TS} \delta_i \Phi \otimes_M \delta_j \Phi \right) + \mathbf{L}_{,vv} :_{\phi_{,M}^* \pi^* E} \left( \nabla^{\phi^* TS} \delta_i \Phi \otimes_M \nabla^{\phi^* TS} \delta_j \Phi \right) \\ &\quad + \mathbf{L}_{,\sigma} \cdot_{\phi_{,M}^* \pi_S^* TS} \nabla \bar{\Phi}_{,M}^{*TS} \partial_i \Phi + \mathbf{L}_{,v} \cdot_{\phi_{,M}^* \pi^* E} \nabla^{\phi^* TS} \nabla \bar{\Phi}_{,M}^{*TS} \partial_i \Phi \\ &\quad + \mathbf{L}_{,v} \cdot_{\phi_{,M}^* \pi^* E} ((\text{Id}_{\phi^* TS} \otimes_M \delta_i \Phi) \cdot_{\phi^* TS \otimes_M \phi^* T^* S} (\phi^* R^{TS} \cdot_{\phi^* TS} \delta_j \Phi) \cdot_{\phi^* TS} \phi_{,M}) \\ &\quad (\text{by Calculation (3)}). \end{aligned}$$

Note that  $\nabla \bar{\Phi}_{,M}^{*TS} \partial_i \Phi \in \Gamma(\phi^* TS)$ , and since  $\phi$  is a critical point of  $\mathcal{L}$ ,

$$\int_M \mathbf{L}_{,\sigma} \cdot_{\phi_{,M}^* \pi_S^* TS} \nabla \bar{\Phi}_{,M}^{*TS} \partial_i \Phi + \mathbf{L}_{,v} \cdot_{\phi_{,M}^* \pi^* E} \nabla^{\phi^* TS} \nabla \bar{\Phi}_{,M}^{*TS} \partial_i \Phi dV_g = 0.$$

Thus

$$\begin{aligned}
& \int_M \nabla^2 (L \circ \Phi_{,M}) :_{TM \oplus TI \oplus TJ} (\delta_i \otimes_{M \times I \times J} \delta_j) dV_g \\
&= \int_M \mathbf{L}_{,\sigma\sigma} \cdot \phi_{,M}^* \pi_S^* TS (\delta_i \Phi \otimes_M \delta_j \Phi) + \mathbf{L}_{,\sigma v} \cdot \phi_{,M}^* \pi_S^* TS \otimes_M \phi_{,M}^* \pi^* E \left( \delta_i \Phi \otimes_M \nabla^{\phi^* TS} \delta_j \Phi \right) \\
&\quad + \mathbf{L}_{,v\sigma} \cdot \phi_{,M}^* \pi^* E \otimes_M \phi_{,M}^* \pi_S^* TS \left( \nabla^{\phi^* TS} \delta_i \Phi \otimes_M \delta_j \Phi \right) + \mathbf{L}_{,vv} \cdot \phi_{,M}^* \pi^* E \left( \nabla^{\phi^* TS} \delta_i \Phi \otimes_M \nabla^{\phi^* TS} \delta_j \Phi \right) \\
&\quad + \delta_i \Phi \cdot \phi^* T^* S \left( \mathbf{L}_{,v} \cdot \phi_{,M}^* \pi^* E \left( (\phi^* R^{TS} \cdot \phi^* TS \delta_j \Phi) \cdot \phi^* TS \phi_{,M} \right) \right) dV_g \\
&= \int_M A \cdot \phi^* T^* S \mathbf{L}_{,\sigma\sigma} \cdot \phi^* TS B + A \cdot \phi^* T^* S \mathbf{L}_{,\sigma v} \cdot \phi^* TS \otimes_M T^* M \nabla^{\phi^* TS} B \\
&\quad + \nabla^{\phi^* TS} A \cdot \phi^* T^* S \otimes_M TM \mathbf{L}_{,v\sigma} \cdot \phi^* TS B + \nabla^{\phi^* TS} A \cdot \phi^* T^* S \otimes_M TM \mathbf{L}_{,vv} \cdot \phi^* TS \otimes_M T^* M \nabla^{\phi^* TS} B \\
&\quad - A \cdot \phi^* T^* S (\mathbf{L}_{,v} \cdot \phi^* TS \otimes_M T^* M (\phi^* R^{TS} \cdot \phi^* TS \phi_{,M})) \cdot \phi^* TS B dV_g \\
&\quad \text{(by antisymmetry of curvature tensor).}
\end{aligned}$$

Calculation (2):

$$\begin{aligned}
& z^* \nabla \nabla \Phi_{,M} :_{z^*(TM \oplus TI \oplus TJ)} (\delta_i \otimes_M \delta_j) \\
&= z^* \nabla \Phi_{,M}^{TE \otimes_M I \times J} (T^* M \oplus T^* I \oplus T^* J) \nabla \Phi_{,M} :_{z^*(TM \oplus TI \oplus TJ)} z^* (\partial_i \otimes_{M \times I \times J} \partial_j) \\
&= z^* \left( \nabla \Phi_{,M}^{TE \otimes_M I \times J} (T^* M \oplus T^* I \oplus T^* J) \nabla \Phi_{,M} :_{TM \oplus TI \oplus TJ} (\partial_i \otimes_{M \times I \times J} \partial_j) \right) \\
&= z^* \left( \nabla_{\partial_j} \Phi_{,M}^{TE \otimes_M I \times J} (T^* M \oplus T^* I \oplus T^* J) \nabla \Phi_{,M} \cdot_{TM \oplus TI \oplus TJ} \partial_i \right) \\
&= z^* \nabla_{\partial_j} \Phi_{,M}^{TE} (\nabla \Phi_{,M} \cdot_{TM \oplus TI \oplus TJ} \partial_i) \quad (\text{since } \nabla_{\partial_j}^{TM \oplus TI \oplus TJ} \partial_i = 0) \\
&= \nabla_{\delta_j} \Phi_{,M}^{TE} \partial_i \Phi_{,M}.
\end{aligned}$$

Calculation (3): As calculated in the proof of (12.1),

$$\begin{aligned}
& \phi_{,M}^* \sigma \cdot \phi_{,M}^* TE \delta_i \Phi_{,M} = \delta_i \Phi \in \Gamma(\phi^* TS), \\
& \phi_{,M}^* \mu \cdot \phi_{,M}^* TE \delta_i \Phi_{,M} = 0 \in \Gamma(TM), \\
& \phi_{,M}^* v \cdot \phi_{,M}^* TE \delta_i \Phi_{,M} = \nabla^{\phi^* TS} \delta_i \Phi \in \Gamma(\phi^* TS \otimes_M T^* M).
\end{aligned}$$

Furthermore, letting  $P := \text{pr}_M^{M \times I \times J}$  for brevity and noting that  $P \circ z = \text{Id}_M$ ,

$$\begin{aligned}
& \phi_{,M}^* \sigma \cdot \phi_{,M}^* TE \nabla_{\delta_j}^{\Phi_{,M}^{TE}} \partial_i \Phi_{,M} \\
&= z^* \nabla_{\partial_j}^{\Phi_{,M}^* \pi_S^* TS} \left( \Phi_{,M}^* \sigma \cdot \Phi_{,M}^* TE \partial_i \Phi_{,M} \right) \quad (\text{since } \nabla \sigma = 0) \\
&= z^* \nabla_{\partial_j}^{(\pi_S \circ \Phi_{,M})^* TS} \partial_i \Phi \quad (\text{using calculation from (12.1)}) \\
&= \nabla_{\delta_j}^{\Phi^* TS} \partial_i \Phi \in \Gamma(z^* \Phi^* TS) \cong \Gamma(\phi^* TS),
\end{aligned}$$

$$\begin{aligned}
& \phi_{*,M}^* \mu \cdot \phi_{*,M}^{*TE} \nabla_{\delta_j}^{\Phi_{*,M}^{*TE}} \partial_i \Phi_{*,M} \\
&= z^* \nabla_{\partial_j}^{\Phi_{*,M}^* \pi_M^{*TM}} \left( \Phi_{*,M}^* \mu \cdot \phi_{*,M}^{*TE} \partial_i \Phi_{*,M} \right) \quad (\text{since } \nabla \mu = 0) \\
&= z^* \nabla_{\partial_j}^{(\pi_M \circ \Phi_{*,M})^* TM} 0 \quad (\text{using calculation from (12.1)}) \\
&= 0 \in \Gamma \left( z^* (\pi_M \circ \Phi_{*,M})^* TM \right) \cong \Gamma \left( z^* P^* TM \right) \cong \Gamma \left( TM \right),
\end{aligned}$$

$$\begin{aligned}
& \phi_{*,M}^* v \cdot \phi_{*,M}^{*TE} \nabla_{\delta_j}^{\Phi_{*,M}^{*TE}} \partial_i \Phi_{*,M} \\
&= z^* \nabla_{\partial_j}^{\Phi_{*,M}^* \pi^{*E}} \left( \Phi_{*,M}^* v \cdot \phi_{*,M}^{*TE} \partial_i \Phi_{*,M} \right) \quad (\text{since } \nabla v = 0) \\
&= z^* \nabla_{\partial_j}^{(\pi \circ \Phi_{*,M})^* E} (\partial_i \Phi)_{*,M} \quad (\text{using calculation from (12.1)}).
\end{aligned}$$

Note that

$$\phi_{*,M}^* v \in \Gamma \left( \phi_{*,M}^* \pi^* E \otimes_M \phi_{*,M}^* T^* E \right) \cong \Gamma \left( (\phi \times_M \text{Id}_M)^* E \otimes_M \phi_{*,M}^* T^* E \right),$$

and therefore

$$\phi_{*,M}^* v \cdot \phi_{*,M}^{*TE} \nabla_{\delta_j}^{\Phi_{*,M}^{*TE}} \partial_i \Phi_{*,M} \in \Gamma \left( (\phi \times_M \text{Id}_M)^* E \right) \cong \Gamma \left( \phi^* TS \otimes_M T^* M \right),$$

so it suffices to examine its natural pairing with  $TM$  elements. Let  $X \in \Gamma(TM)$ , noting that  $X = \text{Id}_M^* X = z^* P^* X$  and that  $P^* X = TP \cdot (X \oplus 0_{TI} \oplus 0_{TJ}) \in \Gamma(P^* TM)$ . Then

$$\begin{aligned}
& \left( \phi_{*,M}^* v \cdot \phi_{*,M}^{*TE} \nabla_{\delta_j}^{\Phi_{*,M}^{*TE}} \partial_i \Phi_{*,M} \right) \cdot_{TM} X \\
&= z^* \nabla_{\partial_j}^{\Phi^* TS \otimes_M \times I \times J P^* T^* M} (\partial_i \Phi)_{*,M} \cdot_{z^* P^* TM} z^* P^* X \\
&= z^* \nabla_{\partial_j}^{\Phi^* TS} \left( (\partial_i \Phi)_{*,M} \cdot_{P^* TM} \nabla P \cdot_{TM \oplus TI \oplus TJ} (X \oplus 0_{TI} \oplus 0_{TJ}) \right) \\
&\quad - z^* \left( (\partial_i \Phi)_{*,M} \cdot \nabla_{\partial_j}^{P^* TM} (\nabla P \cdot_{TM \oplus TI \oplus TJ} (X \oplus 0_{TI} \oplus 0_{TJ})) \right) \\
&= z^* \left( \nabla_{\partial_j}^{\Phi^* TS} \nabla_{X \oplus 0_{TI} \oplus 0_{TJ}}^{\Phi^* TS} \partial_i \Phi - (\partial_i \Phi)_{*,M} \cdot 0_{P^* TM} \right) \\
&= z^* \left( \nabla_{X \oplus 0_{TI} \oplus 0_{TJ}}^{\Phi^* TS} \nabla_{\partial_j}^{\Phi^* TS} \partial_i \Phi + \nabla_{[\partial_j, X \oplus 0_{TI} \oplus 0_{TJ}]}^{\Phi^* TS} \partial_i \Phi - R^{\Phi^* TS} (\partial_j, X \oplus 0_{TI} \oplus 0_{TJ}) \partial_i \Phi \right) \\
&= z^* \left( \nabla_{X \oplus 0_{TI} \oplus 0_{TJ}}^{\Phi^* TS} \nabla_{\partial_j}^{\Phi^* TS} \partial_i \Phi + \nabla_0^{\Phi^* TS} \partial_i \Phi \right) \\
&\quad - z^* \left( (\text{Id}_{\Phi^* TS} \otimes_{M \times I \times J} \partial_i \Phi) \cdot_{\Phi^* TS \otimes_M \times I \times J} R^{\Phi^* TS} :_{TM \oplus TI \oplus TJ} (\partial_j \otimes_{M \times I \times J} (X \oplus 0_{TI} \oplus 0_{TJ})) \right) \\
&= \left[ \nabla_{\delta_j}^{\Phi^* TS} \nabla_{\partial_j}^{\Phi^* TS} \partial_i \Phi + (\text{Id}_{\phi^* TS} \otimes_M \delta_i \Phi) \cdot_{\phi^* TS \otimes_M \phi^* T^* S} (\phi^* R^{TS} \cdot_{\phi^* TS} \delta_j \Phi) \cdot_{\phi^* TS} \phi_{*,M} \right] \cdot_{TM} X,
\end{aligned}$$

where the last equality follows from Calculations (4) and (5). Because  $X$  is pointwise-arbitrary in  $TM$ , this shows that

$$\phi_{*,M}^* v \cdot \phi_{*,M}^{*TE} \nabla_{\delta_j}^{\Phi_{*,M}^{*TE}} \partial_i \Phi_{*,M} = \left( \nabla_{\delta_j}^{\Phi^* TS} \partial_i \Phi \right)_{*,M} + (\text{Id}_{\phi^* TS} \otimes_M \delta_i \Phi) \cdot_{\phi^* TS \otimes_M \phi^* T^* S} (\phi^* R^{TS} \cdot_{\phi^* TS} \delta_j \Phi) \cdot_{\phi^* TS} \phi_{*,M}.$$

Calculation (4):

$$\begin{aligned}
& z^* \left( \nabla_{X \oplus 0_{TI} \oplus 0_{TJ}}^{\Phi^* TS} \nabla_{\partial_j}^{\Phi^* TS} \partial_i \Phi + \nabla_0^{\Phi^* TS} \partial_i \Phi \right) \\
&= z^* \left( \nabla_{\partial_j}^{\Phi^* TS} \partial_i \Phi \right)_{*,M} \cdot_{z^* P^* TM} z^* P^* X \\
&= \left( \nabla_{\delta_j}^{\Phi^* TS} \partial_i \Phi \right)_{*,M} \cdot_{TM} X \quad (\text{by (10.6)}) \\
&= \nabla_{\delta_j}^{\Phi^* TS} \nabla_{\partial_j}^{\Phi^* TS} \partial_i \Phi \cdot_{TM} X \quad (\text{because } \nabla_{\delta_j}^{\Phi^* TS} \partial_i \Phi \in \Gamma(z^* \Phi^* TS) \cong \Gamma(\phi^* TS)).
\end{aligned}$$



Calculation (5):

$$\begin{aligned}
& -z^* \left( R^{\Phi^*TS} :_{TM \oplus TI \oplus TJ} (\partial_j \otimes_{M \times I \times J} (X \oplus 0_{TI} \oplus 0_{TJ})) \right) \\
&= z^* \left( R^{\Phi^*TS} :_{TM \oplus TI \oplus TJ} ((X \oplus 0_{TI} \oplus 0_{TJ}) \otimes_{M \times I \times J} \partial_j) \right) \quad (\text{antisymmetry of } R^{\Phi^*TS}) \\
&= z^* \left( \Phi^* R^{TS} :_{\Phi^*TS} (\nabla \Phi \boxtimes_{M \times I \times J} \nabla \Phi) :_{TM \oplus TI \oplus TJ} ((X \oplus 0_{TI} \oplus 0_{TJ}) \otimes_{M \times I \times J} \partial_j) \right) \quad (\text{by (10.5)}) \\
&= z^* \left( \Phi^* R^{TS} :_{\Phi^*TS} ((\Phi_{,M} \cdot P^*TM P^*X) \otimes_{M \times I \times J} \partial_j \Phi) \right) \\
&= z^* \left( (\Phi^* R^{TS} \cdot_{\Phi^*TS} \partial_j \Phi) \cdot_{\Phi^*TS} \Phi_{,M} \cdot P^*TM P^*X \right) \\
&= (z^* \Phi^* R^{TS} \cdot_{z^* \Phi^*TS} z^* \partial_j \Phi) \cdot_{z^* \Phi^*TS} z^* \Phi_{,M} \cdot_{z^* P^*TM} z^* P^*X \\
&= (\phi^* R^{TS} \cdot_{\phi^*TS} \delta_j \Phi) \cdot_{\phi^*TS} \phi_{,M} \cdot_{TM} X.
\end{aligned}$$

□

**Theorem 13.2** (Second variation of  $\mathcal{L}$  (alternate form)). *Let  $\mathcal{L}$ ,  $L$ ,  $\sigma$ ,  $\mu$ ,  $\nu$  and  $\nu$  all be defined as above. If  $\phi \in C^\infty(M, S)$  is a critical point of  $\mathcal{L}$  and  $A, B \in \Gamma(\phi^*TS)$ , then*

$$\begin{aligned}
& \nabla^2 \mathcal{L}(\phi) :_{T_\phi C} (A \otimes B) \\
&= \int_M A \cdot_{\phi^*T^*S} \mathbf{L}_{,\sigma\sigma} \cdot_{\phi^*TS} B + A \cdot_{\phi^*T^*S} \mathbf{L}_{,\sigma\nu} \cdot_{\phi^*TS \otimes_M T^*M} \nabla^{\phi^*TS} B \\
&\quad - A \cdot_{\phi^*T^*S} \text{div}_M \mathbf{L}_{,\nu\sigma} \cdot_{\phi^*TS} B - A \cdot_{\phi^*T^*S} \mathbf{L}_{,\nu\sigma} \cdot_{T^*M \otimes_M \phi^*TS} \left( \nabla^{\phi^*TS} B \right)^{(12)} \\
&\quad - A \cdot_{\phi^*T^*S} \text{div}_M \mathbf{L}_{,\nu\nu} \cdot_{\phi^*TS \otimes_M T^*M} \nabla^{\phi^*TS} B \\
&\quad - A \cdot_{\phi^*T^*S} \mathbf{L}_{,\nu\nu} \cdot_{T^*M \otimes_M \phi^*TS \otimes_M T^*M} \left( \nabla^{\phi^*TS \otimes_M T^*M} \nabla^{\phi^*TS} B \right)^{(123)} \\
&\quad - A \cdot_{\phi^*T^*S} (\mathbf{L}_{,\nu} \cdot_{\phi^*TS \otimes_M T^*M} (\phi^* R^{TS} \cdot_{\phi^*TS} \phi_{,M})) \cdot_{\phi^*TS} B dV_g \\
&\quad + \int_{\partial M} (A \cdot_{\phi^*T^*S} \mathbf{L}_{,\nu\sigma} \cdot_{\phi^*TS} B) \cdot_{T^*M} \nu + \left( A \cdot_{\phi^*T^*S} \mathbf{L}_{,\nu\nu} \cdot_{\phi^*TS \otimes_M T^*M} \nabla^{\phi^*TS} B \right) \cdot_{T^*M} \nu d\bar{V}_g
\end{aligned}$$

*Proof.* This result follows essentially from (13.1) via several instances of integration by parts to express the integrand(s) entirely in terms of  $A$  and not its covariant derivatives. Abbreviate  $\phi_{,M}^* L_{,xy}$  by  $\mathbf{L}_{,xy}$ . Then, integrating by parts allows the covariant derivatives of  $A$  to be flipped across the natural pairings over  $\phi^*TS$ .

$$\begin{aligned}
& \int_M \nabla^{\phi^*TS} A \cdot_{\phi^*T^*S \otimes_M TM} \mathbf{L}_{,\nu\sigma} \cdot_{\phi^*TS} B dV_g \\
&= \int_M \text{tr}_{TM} \left( \left( \nabla^{\phi^*TS} A \right)^{(12)} \cdot_{\phi^*TS} \mathbf{L}_{,\nu\sigma} \cdot_{\phi^*TS} B \right) dV_g \quad (TM\text{trace is taken separately}) \\
&= \int_M \text{tr}_{TM} (\nabla^{TM} (A \cdot_{\phi^*T^*S} \mathbf{L}_{,\nu\sigma} \cdot_{\phi^*TS} B)) \\
&\quad - \text{tr}_{TM} \left( A \cdot_{\phi^*T^*S} \nabla^{\phi^*T^*S \otimes_M TM \otimes_M \phi^*TS} \mathbf{L}_{,\nu\sigma} \cdot_{\phi^*TS} B \right) \\
&\quad - \text{tr}_{TM} \left( A \cdot_{\phi^*T^*S} \mathbf{L}_{,\nu\sigma} \cdot_{\phi^*TS} \nabla^{\phi^*TS} B \right) dV_g \quad (\text{reverse product rule}) \\
&= \int_M -A \cdot_{\phi^*T^*S} \text{div}_M \mathbf{L}_{,\nu\sigma} \cdot_{\phi^*TS} B \quad (\text{definition of divergence}) \\
&\quad - A \cdot_{\phi^*T^*S} \mathbf{L}_{,\nu\sigma} \cdot_{T^*M \otimes_M \phi^*TS} \left( \nabla^{\phi^*TS} B \right)^{(12)} dV_g \\
&\quad + \int_{\partial M} (A \cdot_{\phi^*T^*S} \mathbf{L}_{,\nu\sigma} \cdot_{\phi^*TS} B) \cdot_{T^*M} \nu d\bar{V}_g \quad (\text{divergence theorem}).
\end{aligned}$$

Similiarly,

$$\begin{aligned}
& \int_M \nabla^{\phi^* TS} A \cdot_{\phi^* T^* S \otimes_M TM} \mathbf{L}_{,vv} \cdot_{\phi^* TS \otimes_M T^* M} \nabla^{\phi^* TS} B dV_g \\
&= \int_M \text{tr}_{TM} \left( \left( \nabla^{\phi^* TS} A \right)^{(12)} \cdot_{\phi^* T^* S} \mathbf{L}_{,vv} \cdot_{\phi^* TS \otimes_M T^* M} \nabla^{\phi^* TS} B \right) dV_g \\
&= \int_M \text{tr}_{TM} \left( \nabla^{TM} \left( A \cdot_{\phi^* T^* S} \mathbf{L}_{,vv} \cdot_{\phi^* TS \otimes_M T^* M} \nabla^{\phi^* TS} B \right) \right) \\
&\quad - \text{tr}_{TM} \left( A \cdot_{\phi^* T^* S} \nabla^{\phi^* T^* S \otimes_M TM \otimes_M \phi^* T^* S \otimes_M TM} \mathbf{L}_{,vv} \cdot_{\phi^* TS \otimes_M T^* M} \nabla^{\phi^* TS} B \right) \\
&\quad - \text{tr}_{TM} \left( A \cdot_{\phi^* T^* S} \mathbf{L}_{,vv} \cdot_{\phi^* TS \otimes_M T^* M} \nabla^{\phi^* TS \otimes_M T^* M} \nabla^{\phi^* TS} B \right) dV_g \\
&= \int_M -A \cdot_{\phi^* T^* S} \text{div}_M \mathbf{L}_{,vv} \cdot_{\phi^* TS \otimes_M T^* M} \nabla^{\phi^* TS} B \\
&\quad - A \cdot_{\phi^* T^* S} \mathbf{L}_{,vv} \cdot_{T^* M \otimes_M \phi^* TS \otimes_M T^* M} \left( \nabla^{\phi^* TS \otimes_M T^* M} \nabla^{\phi^* TS} B \right)^{(123)} dV_g \\
&\quad + \int_{\partial M} \left( A \cdot_{\phi^* T^* S} \mathbf{L}_{,vv} \cdot_{\phi^* TS \otimes_M T^* M} \nabla^{\phi^* TS} B \right) \cdot_{T^* M} \nu d\bar{V}_g.
\end{aligned}$$

Together with (13.1), this gives the desired result.  $\square$

## 14 Questions and Future Work

This paper is a first pass at the development of a strongly-typed tensor calculus formalism. The details of its workings are by no means complete or fully polished, and its landscape is riddled with many tempting rabbit holes which would certainly produce useful results upon exploration, but which were out of the scope of a first exposition. Here is a list of some topics which the author considers worthwhile to pursue, and which will likely be the subject of his future work. Hopefully some of these topics will be inspiring to other mathematicians, and ideally will start a conversation on the subject.

- There refinements to be made to the type system used in this paper in order to achieve better error-checking and possibly more insight into the relevant objects. There are still implicit type identifications being done (mostly the canonical identifications between different pullback bundles).
- The calculations done in this paper are not in an optimally polished and refined state. With experience, certain common operations can be identified, abstract computational rules generated for these operations, and the relevant calculations simplified.
- The language of Category Theory can be used to address the implicit/explicit handling of natural type identifications, for example, the identification used in showing the contravariance of bundle pullback;  $\psi^* \phi^* F \cong (\phi \circ \psi)^* F$ .
- The details of the particular implementation of the pullback bundle  $\phi^* F$  as a submanifold of the direct product  $M \times F$  are used in this paper, but there is no reason to “open up the box” like this. For most purposes, the categorical definition of pullback bundle suffices; the pullback bundle can be worked exclusively using its projection maps  $\pi_M^{\phi^* F}$  and  $\rho_F^{\phi^* F}$ . In the author’s experience (which occurred too late to be incorporated into this paper), using this abstract interface cleans up calculations involving pullback bundles significantly.
- The type system used for any particular problem or calculation can be enriched or simplified to adjust to the level of detail appropriate for the situation. For example, if  $\gamma \in C^\infty(\mathbb{R}, M)$ , then  $\nabla \gamma \in \Gamma(\gamma^* TM \otimes_{\mathbb{R}} T^* \mathbb{R})$ , but if  $t$  is the standard coordinate on  $\mathbb{R}$ , then  $\nabla \gamma = \gamma' \otimes_{\mathbb{R}} dt$ , where  $\gamma' \in \Gamma(\gamma^* TM)$  is given by  $\nabla \gamma \cdot \frac{d}{dt}$ . This “primed” derivative has a simpler type than the total derivative, and would

presumably lead to simpler calculations (e.g. in (12.6)). This “primed” derivative could also be used in the derivation of the first and second variations. While this would simplify the type system, it would diversify the notation and make the computational system less regularized. However, some situations may benefit overall from this.

- The notion of strong typing comes from computer programming languages. The human-driven type-checking which is facilitated by the pedantically decorated notation in this paper can be done by computer by implementing the objects and operations of this tensor calculus formalism in a strongly typed language such as Haskell. This would be a step toward automated calculation checking, and could be considered a step toward automated proof checking from the top down (as opposed to from the bottom up, using a system such as the Coq Proof Assistant).
- Is there some sort of completeness result about the calculational tools and type system in this paper? In other words, is it possible to accomplish “everything” in a global, coordinate-free way using a certain set of tools, such as pullback bundles, covariant derivatives, chain rules, permutations, evaluation-by-pullback?
- The alternate form of the second variation (see (13.2)) can be used to form a generalized Jacobi field equation for a particular energy functional. Analysis of this equation and its solutions may give insights analogous to the standard (geodesic-based) Jacobi field equation.

## Acknowledgements

I would like to express my gratitude to the ARCS (Achievement Rewards for College Scientists) Foundation for their having awarded me a 2011-2012 ARCS Fellowship, and for their generous efforts to promote excellence in young scientists. I would like to thank my advisor Debra Lewis for trusting in my abilities and providing me with the freedom in which the creative endeavor that this paper required could flourish. I would like to thank David DeConde for the invaluable conversations at the Octagon in which imagination, creativity, and exploration were gladly fostered. Thanks to Chris Shelley for showing me how to create the tensor diagrams using Tikz. Finally, I would like to thank both Debra and David for their help in editing this paper.

## References

- [1] Luca Cardelli. Typeful programming. 1991 (revised 1993). Available online at [ftp://gatekeeper.research.compaq.com/pub/DEC/SRC/research-reports/SRC-045.pdf](http://gatekeeper.research.compaq.com/pub/DEC/SRC/research-reports/SRC-045.pdf).
- [2] Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics. *Selected Topics in Harmonic Maps*, number 50. American Mathematical Society, 1983.
- [3] C.T.J. Dodson and M.S. Radivoiović. Second-order tangent structures. *International Journal of Theoretical Physics*, 21(2):151–161, 1982.
- [4] David G. Ebin and Jerrold E. Marsden. Groups of diffeomorphisms and the motion of an incompressible fluid. *The Annals of Mathematics, Second Series*, 92(1):102–163, 1970.
- [5] Halldor I. Eliasson. Geometry of manifolds of maps. *J. Differential Geometry*, 1(2), 1967.
- [6] Mariano Giaquinta and Stefan Hildebrandt. *Calculus of Variations I*. Springer-Verlag, 1996.
- [7] Ivan Kolár, Peter W. Michor, and Jan Slovák. *Natural Operations in Differential Geometry*, volume 434. Springer Verlag, 1993. This is an online book which can be found at <http://www.mat.univie.ac.at/michor/listpubl.html>.

- [8] Jeffrey M. Lee. *Manifolds and Differential Geometry*, volume 107. American Mathematical Society, 2009.
- [9] John M. Lee. *Riemannian Manifolds: An Introduction to Curvature*, volume 176. Springer Verlag, 1997.
- [10] John M. Lee. *Introduction to Smooth Manifolds*, volume 218. Springer Verlag, 2006.
- [11] Jerrold E. Marsden and Thomas J. R. Hughes. *Mathematical Foundations of Elasticity*. Prentice Hall, Inc., 1983.
- [12] Peter W. Michor. *Topics in Differential Geometry*, volume 93. American Mathematical Society, 2008.
- [13] George A. Miller. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *American Psychological Association*, 101(2):343–352, 1955.
- [14] Seiki Nishikawa. *Variational Problems in Geometry*, volume 205. American Mathematical Society, 2002.
- [15] Richard S. Palais. *Foundations of Global Non-Linear Analysis*. W.A. Benjamin, Inc., 1968.
- [16] David Parnas. On the criteria to be used in decomposing systems into modules. *Communications of the ACM*, 15(12):1053–1058, 1972.
- [17] Roger Penrose. *The Road to Reality*. Vintage Books, 2004.
- [18] Eric S. Raymond. *The Art of Unix Programming*. Pearson Education, Inc., 2003. This is an online book which can be found at <http://www.faqs.org/docs/artu/index.html>.
- [19] Wolfgang Walter. *Ordinary Differential Equations*, volume 182. Springer Verlag, 1998.
- [20] Yuanlong Xin. *Geometry of Harmonic Maps*, volume 23. Birkhäuser, 1996.